

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2001-135692
(P2001-135692A)

(43) 公開日 平成13年5月18日 (2001.5.18)

(51) Int.Cl. ⁷	識別記号	F I	テ-マコ-ト* (参考)
H 0 1 L 21/66		H 0 1 L 21/66	J
G 0 1 B 11/30		G 0 1 B 11/30	A
G 0 1 N 21/956		G 0 1 N 21/956	A

審査請求 未請求 請求項の数30 O L 外国語出願 (全 40 頁)

(21) 出願番号 特願2000-219833(P2000-219833)
(22) 出願日 平成12年6月15日 (2000.6.15)
(31) 優先権主張番号 09/334049
(32) 優先日 平成11年6月15日 (1999.6.15)
(33) 優先権主張国 米国 (US)

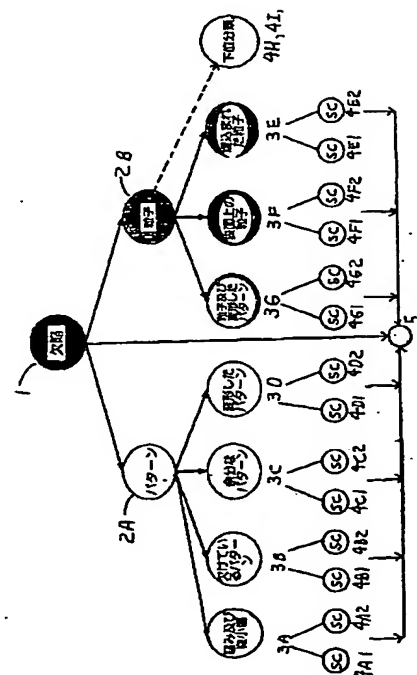
(71) 出願人 390040660
アプライド マテリアルズ インコーポレ
イテッド
APPLIED MATERIALS, I
NCORPORATED
アメリカ合衆国 カリフォルニア州
95054 サンタ クララ パウアーズ ア
ベニュー 3050
(72) 発明者 アリエル ペン ポラト
イスラエル レボヴォト ピンスカー ス
トリート14
(74) 代理人 100059959
弁理士 中村 裕 (外9名)

(54) 【発明の名称】 ハイブリッドで一様に適用可能な自動欠陥分類法

(57) 【要約】

【課題】 自動的、迅速、確実及び完全な半導体ウェーハの欠陥の分類のための方法及び装置を供給する。

【解決手段】 方法及び装置は、境界及び地形情報を用いるコア分級機を使用して半導体ウェーハの表面の欠陥をコア分類の所定数の1つに自動的に分類するための提供されている。その後、欠陥は、任意に定義された欠陥の下位分類に分類され、1つのコア分類と結合され、関連するコア分類の限定数からのみ欠陥を分類するように調整された特定の適用可能な分級機でユーザにより定義される。コア分級機又は特定の適用可能な分級機によって分類されることができない欠陥は完全な分級機により分類される。



【特許請求の範囲】

【請求項1】 物品の表面の欠陥を分類する方法であって、
前記表面を映し、欠陥画像を形成し、
欠陥のコア分類の所定数の1つであるとして、コア分級機を使用して前記欠陥を分類し、
別の下位分類の任意数の1つであるとして、前記1つのコア分類に関連した特定の適用可能な分級機を使用して前記欠陥を分類する、ことを含むことを特徴とする方法。

【請求項2】 コアの分級機がコアの分類の1つに前記欠陥を分類できない時、別の分類の任意数の1つであるとして、完全な分級機を使用して前記欠陥を分類することを含ま請求項1に記載の方法。

【請求項3】 欠陥の前記コア分類はパターンの欠陥と粒子の欠陥を含んでいる請求項1に記載の方法。

【請求項4】 欠陥の前記コア分類は前記表面上の欠けているパターン、前記表面上の余分なパターン、前記表面上の粒子、前記表面に埋め込まれた粒子、及び前記表面上の微小傷を含んでいる請求項3に記載の方法。

【請求項5】 欠陥の前記別の下位分類は前記表面上の、又は前記表面に埋め込まれた所定寸法の粒子を含んでいる請求項4に記載の方法。

【請求項6】 複数の特定の適用可能な分級機を供給することを含み、それぞれはコア分類の所定数以下と関連している請求項1に記載の方法。

【請求項7】 前記欠陥画像の境界分析、前記欠陥画像の地形分析、又は前記欠陥を前記1つのコア分類に分類する両方を行なうことを含んでいる請求項1に記載の方法。

【請求項8】 前記特定の適用可能な分級機を訓練し、前記別の下位分類の欠陥の公知な属性の例示的なセットを基に前記別の下位分類の欠陥を確認し、
前記特定の適用可能な分級機を使用して前記欠陥画像に関連する属性を分析し、前記別の下位分類の1つに前記欠陥を分類する、ことを含んでいる請求項7に記載の方法。

【請求項9】 判断ツリー又は多次元のクラスタ技術を使用して前記特定の適用可能な分級機を訓練することを含んでいる請求項8に記載の方法。

【請求項10】 物品の前記表面の欠陥を分類するための装置であって、
前記欠陥の画像を作る画像装置と、
欠陥のコア分類の所定数の1つであるとして前記欠陥を分類するためのコア分級機と、別の下位分類の任意数の1つであるとして前記欠陥を分類するための前記1つのコア分類と関連した特定の適用可能な分級機とを備えたプロセッサと、を備えたことを特徴とする装置。

【請求項11】 前記プロセッサは、前記コア分級機が前記コア分類の1つに前記欠陥を分類することができな

い時、別の分類の任意数の1つであるとして前記欠陥を分類するための完全な分級機をさらに備えている請求項10に記載の装置。

【請求項12】 欠陥の前記コア分類はパターン上の欠陥又は粒子の欠陥を含んでいる請求項10に記載の装置。

【請求項13】 欠陥の前記コア分類は前記表面上の欠けているパターン、前記表面上の余分なパターン、前記表面上の粒子、前記表面に埋め込まれた粒子、及び前記表面上の微小傷を含んでいる請求項12に記載の装置。

【請求項14】 欠陥の前記別の下位分類は前記表面上の、又は前記表面に埋め込まれた所定寸法の粒子を含んでいる請求項13に記載の装置。

【請求項15】 複数の特定の適用可能な分級機を供給することを含み、それぞれはコア分類の所定数以下と関連している請求項10に記載の装置。

【請求項16】 前記プロセッサはさらに、前記欠陥画像の境界分析、前記欠陥画像の地形分析、又は前記欠陥を前記1つのコア分類に分類する両方を行なうように構成されている請求項10に記載の装置。

【請求項17】 前記特定の適用可能な分級機は、前記別の下位分類の欠陥の公知な属性の例示的なセットを基に前記別の下位分類の欠陥を確認するように訓練可能であり、

前記特定の適用可能な分級機は、前記欠陥画像に関連する属性を分析し、前記別の下位分類の1つに前記欠陥を分類するためである、請求項16に記載の装置。

【請求項18】 前記特定の適用可能な分級機は、判断ツリー又は多次元のクラスタ技術を使用して訓練可能である請求項17に記載の装置。

【請求項19】 前記コア分類にあるとして前記欠陥を分類するコア分級機に反応して、コア欠陥分類の別の下位分類の任意数の1つであるとして物品の前記表面の欠陥を分類することを特徴とする特定の適用可能な分級機。

【請求項20】 前記特定の適用可能な分級機は前記別の下位分類の欠陥の公知な属性の例示的なセットを基に前記別の下位分類の欠陥を確認するように訓練可能であり、
前記特定の適用可能な分級機は、前記欠陥画像と関連する属性を分析し、前記別の下位分類の1つに前記欠陥を分類するためである、請求項19に記載の特定の適用可能な分級機。

【請求項21】 前記特定の適用可能な分級機は判断ツリー又は多次元のクラスタ技術を使用して訓練可能である請求項20に記載の特定の適用可能な分級機。

【請求項22】 物品の表面上の欠陥を自動的に分類するための命令を有するコンピュータで読取り可能な媒体であって、実行する時に、前記命令は、
前記表面を映し、欠陥画像を形成し、
欠陥のコア分類の所定数の1つであるとして前記欠陥を

分類し、

前記1つのコア分類であるとして前記欠陥の分類を基に別の下位分類の任意数の1つであるとして前記欠陥を分類する、

段階を1以上のプロセッサに行わせるように配列されていることを特徴とするコンピュータで読取り可能な媒体。

【請求項23】 実行する時に、前記命令は、前記1以上のプロセッサが前記欠陥を前記コア分類の1つに分類することができない時、別の分類の任意数の1つであるとして前記欠陥を分類する段階を前記1以上のプロセッサに行わせるように配列されている請求項22に記載のコンピュータで読取り可能な媒体。

【請求項24】 欠陥の前記コア分類はパターンの欠陥と粒子の欠陥を備えている請求項22に記載のコンピュータで読取り可能な媒体。

【請求項25】 欠陥の前記コア分類は前記表面上の欠けているパターン、前記表面上の余分なパターン、前記表面上の粒子、前記表面に埋め込まれた粒子、及び前記表面上の微小傷を含んでいる請求項24に記載のコンピュータで読取り可能な媒体。

【請求項26】 欠陥の前記別の下位分類は前記表面上の、又は前記表面に埋め込まれた所定寸法の粒子を含んでいる請求項25に記載のコンピュータで読取り可能な媒体。

【請求項27】 実行する時に、前記命令は、前記欠陥画像の境界分析、前記欠陥画像の地形分析、又は前記欠陥を前記1つのコア分析に分析する両方を前記1以上のプロセッサに行なわせるように配列されている請求項22に記載のコンピュータで読取り可能な媒体。

【請求項28】 実行する時に、前記命令は、前記別の下位分類の欠陥の公知な属性の例示的セットを基に前記別の下位分類の欠陥を確認するように学習し、前記欠陥画像と関連する属性を分析し、前記別の下位分類の1つに前記欠陥を分類する、段階を前記1以上のプロセッサに行なわせるように配列された請求項26に記載のコンピュータで読取り可能な媒体。

【請求項29】 実行する時に、前記命令は、判断ツリー又は多次元のクラスタ技術を使用して前記学習段階を前記1以上のプロセッサに行なわせるように配列されている請求項28に記載のコンピュータで読取り可能な媒体。

【請求項30】 実行する時に、前記命令は、走査型電子顕微鏡で前記表面を映す段階を前記1以上のプロセッサに行なわせるように配列されている請求項22に記載のコンピュータで読取り可能な媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、物品の表面の欠陥を自動的に分類するための方法及び装置に関する。本発

明はミクロン以下の設計特徴の高密度の半導体デバイスの製造の間に半導体ウェーハのインライン検査のために特に適用するものである。

【0002】

【従来の技術】超大規模集積に伴う高密度及び高性能の現在の要求は、ミクロン以下の外形、増加したトランジスタ及び回路速度及び改善した信頼性を要求する。そのような要求は高精度及び均一性のデバイスの特徴の形成を要求し、順番に、注意深い処理の監視を必要とし、それらがまだ半導体ウェーハの形にある間、デバイスの頻繁で詳細な検査を備えている。

【0003】伝統的な製造過程の監視技術は検査及び再調査手続きを使用し、ウェーハの表面は最初、高速度で比較的低い解像度のツール、例えば、CCD（電荷結合素子）又はレーザのような光・電気変換器により走査される。その後、統計的方法が使用され、欠陥のおそれの高いウェーハ上に疑わしい位置を示す検出マップを作り出す。潜在的な欠陥の数及び又は密度が所定レベルに達した場合には、警報が鳴らされ、潜在的な欠陥の場所のより詳細な様子が保証されることを示している。この技術は欠陥の「全体の密度監視」として公知であり、「全体の欠陥密度」と呼ばれる統計値を作り出す。

【0004】欠陥密度が所定レベルに達すると、影響を受けたウェーハの再調査が保証される。検出マップを使用して、再検出手続きが行なわれ、欠陥の存在を明確に決定した後、より詳細な再調査手続きがCCDで走査するような個々の欠陥場所で行われ、比較的高解像度の画像を作り出す。その後、欠陥画像が分析され、欠陥の性質（例えば、欠陥パターン、粒子、又は傷）を決定する。

【0005】最近、分類された欠陥密度、すなわち、幾つかの別のタイプの欠陥数、又は欠陥の分類を監視することが認められているので、欠陥を分類するための各種方法が紹介されている。ここでは「古典的分級機」と呼ばれるこれらの伝統的な方法のほとんどはパターン認識技術を使用し、1セットのサンプル欠陥が特性又は属性（例えば、輝度、粗さ、寸法、色）のため得られ、映され、そして分析され、属性は「ブラックボックス」（例えば、ニューラルネット）に与えられ、分級機を訓練するために使用され、欠陥の属性により欠陥の別のタイプを認識する。

【0006】不都合にも、同意されたセットの欠陥分類がないので、これらの方法の効率は減少される。別の半導体製作者は別の欠陥を重要であると考え、そのため、それらの特定のニーズに合わせて作られた別のセットの欠陥分類を使用する。したがって、それらは作動される前に各欠陥分類のために得られる多くの例の欠陥画像を要求する。その結果として、通常の従来の技術のシステムは製造ラインの開始及び取り決めの間、使用されることはできない。その上さらに、ここでは「完全な分級

機」とも呼ばれているそのような分級機は分類されるために要求されるすべての欠陥タイプ（例えば、10以上の欠陥分類）の間で識別されることが必要であるので、欠陥を分類する時に非常に多くの属性が考慮されなければならない、したがって、検査時間を増加し、製造スループットを減少させる。

【0007】完全な分級機と関連した上述した問題点と取り組むため、一様なコア分級機（ここでは「コア分級機」）が最近、欠陥再調査システムで導入され、それはカリフォルニア州のサンタクララ市の Applied Materials社より入手可能なSEMVision（登録商標）として注目されている。そのようなコア分級機は1998年7月8日に出願された「一様なコア分類を有する自動欠陥分類（Automatic Defect Classification）」という表題の共同出願中の米国特許出願No. 09/111,454で説明されており、その完全な開示がここにインコーポレイテッドバイリファレンスされている。

【0008】その共同出願中の出願の方法論によれば、半導体ウェーハの欠陥マップが生成された後、各欠陥場所及び相当する公知の欠陥のない参照場所は走査電子顕微鏡（SEM）により映され、位置及び地形学データを集めると共に蓄積する。その後、境界分析及び又は地形学データを行なうことにより画像が分析され、多数（例えば、7個）の一様な欠陥コア分類の1つであるとして欠陥を分類し、コア分級機に予めプログラムされたモジュールを付加することにより、ユーザーにより所望のように任意数のコアの下位分類の1つであるとしてさらに分類される。

【0009】図1は共同出願中の出願の地形学により行われたコア分類への自動欠陥分類の概念フローチャートである。欠陥1はパターン欠陥2A又は粒子欠陥2Bとして広く分類され、ウェーハ表面の窪み及び微小傷3A、表面の欠けているパターン3B、表面の余分なパターン3C、表面の変形パターン3D、表面の粒子3E、表面に埋め込まれた粒子3F、又は表面の粒子及び変形パターン3Gという7個の例示の一様な欠陥のコア分類の1つにさらに配置される。任意のコア下位分類は隣接の配線パターン、小粒子、大粒子、破線、狭いパターン等の間のブリッジング（すなわち、短絡）を含んでいてもよい。

【0010】共同出願の出願の一様なコア分類技術は欠陥を、粒子又はパターン欠陥として、及び表面上又は表面下（埋め込まれた）欠陥として別々及び確実に分類させる。それはまた、これらの意味を持った分類された欠陥の早い定量化及び通知を供給し、それにより、欠陥の原因の調査を容易にし、実行される早い修正動作を可能にする。

【0011】共同出願中の出願のコア分級機はそれが欠陥情報を収集（すなわち、ウェーハ表面を映し、境界分析及び又はその外形の地形学測定を実行）することによ

り欠陥を分類し、その後、プリオリ（priori）（すなわち、事前に）プログラムされた1セットのルールに続く点で、ルールベースの分級機である。したがって、伝統的な分級機をするようにそれは訓練される必要はなく、それで、作動する前に各分類のための欠陥画像の例を必要としない。そのため、従来技術の欠陥分類システムとは異なり、共同出願中の出願のコア分級機は製造ラインの開始及び取り決めの間に使用可能である。

【0012】共同出願中の出願で説明されているコア分級機は伝統的な分級機の多くの欠点に取り組んでいるが、ルールベースの分級機として、それらは新しい欠陥分類を認めるように容易に適応することができないので、コア分級機はユーザーが重要だと考えたすべての分類に欠陥を分類するのに適していなくてもよい。特に、（異なる処理ラインはお互いとは別の欠陥に敏感であってもよいので）ユーザーはコア分級機に付加される予めプログラムされたモジュールとして利用可能な上述したコアの下位分類以外の一様なコア分類内の改善を要求してもよい。その上さらに、ユーザーはコア分級機により認識できない改善を要求してもよい。例えば、コア分類が表面の粒子として欠陥を分類し（図1のコア分類3F）、ユーザーがそのサイズと共に粒子の形状を知りたい場合には、別の処理が別の粒子形状及び寸法を作る傾向があるので、別の技術はこの寸法情報を得るために使用されなければならない、それは粒子の源を正確に示すのに役立つ。さらに、コア分類のいずれにも分類されない新型の欠陥はコア分級機により分類することができない。例えば、処理が導入され、新しいタイプの欠陥となった場合には、現存のコア分類は新しい欠陥と関連して不適切であり、コア分級機は新しい欠陥がコア分類として付加されなければ新しい欠陥を分類することができないであろう。

【0013】コア分級機を使用して、半導体ウェーハを迅速に意義のある再調査をすると共に欠陥を自動的に分類し、その後、欠陥の原因となる処理を確認するためユーザーにより望まれるコア分類内の下位分類に欠陥をさらに分類し、それにより、早い修正動作をとらせる必要性がある。この必要性は表面外形の密度、ダイの寸法、及びデバイスの層数が増加するにつれて、より重要となり、欠陥数を徹底的に減少させ、許容の製造歩留りを達成する。

【0014】

【発明が解決しようとする課題】本発明の目的は、自動的、迅速、確実及び完全な半導体ウェーハの欠陥の分類のための方法及び装置を供給することである。

【0015】

【課題を解決するための手段】本発明によれば、前述及び他の目的は、物品の表面の欠陥を分類する方法により部分的に達成され、その方法は表面を映し、コア分級機を使用して欠陥のコア分類の所定数の1つとして欠陥を

分類し、1つのコア分類と関連して特定の適応可能な分級機を使用して一様な下位分類の任意数の1つとして欠陥を分類することを含んでいる。コア分級機が欠陥にコア分類を分類することができない場合、完全な分級機が使用され、欠陥を分類する。

【0016】本発明のさらなる特徴は、部品野表面の欠陥を自動的に分類するための命令を伝えるコンピュータで読取り可能な媒体であり、その命令は実行される時、1以上のプロセッサに上記方法の段階を行なわせるように配置されている。

【0017】本発明のさらなる目的及び利点は当業者であれば以下の詳細な説明から容易に明らかであり、本発明を実行するために考えられたベストモードの実例により簡単に本発明の好適な実施例のみが示されると共に説明されている。分かるように、本発明は他の異なる実施例が可能であり、その幾つかの詳細は本発明から逸脱することなく各種明らかな点での変更が可能である。したがって、図面及び説明は事実上、制限ではなく例示とみなされるべきである。

【0018】

【発明の実施の形態】共同出願中の米国出願09/111,454等のコア分級機は欠陥を、粒子又はパターン欠陥として、及び表面上又は表面下（埋め込まれた）欠陥として、迅速且つ確実に分類させ、製造ラインの開始及び取り決めの間、動作可能である。しかし、コア分級機はそのコア分類内の一定のタイプの欠陥を分類することができず、欠陥分類を付加するためにユーザにより容易に変更されない。一方、伝統的な分級機はユーザにより訓練されるような如何なるタイプの欠陥も完全に分類するようにすることができるが、作動される前に各欠陥分類のため得られる多くの例の欠陥画像を必要とする。したがって、伝統的な分級機は製造ラインの開始及び取り決めの間、使用することができない。その上さらに、伝統的なタイプの完全な分級機はコア分級機より遅く、通常、あまり正確ではない。本発明はハイブリッドな欠陥分類方法学を供給することによりこれらの問題に取り組むと共に解決し、ユーザにより必要とされるような一様なコア分類及び別の下位分類への欠陥の迅速且つ確実な分類のため限定範囲の伝統的なタイプの分級機と共にコア分級機を利用する。したがって、現在の方法学は欠陥を引き起こす処理を容易に確認でき、それにより、採るべき素早い修正動作を可能とする。

【0019】本発明の方法学のある実施例によると、半導体ウェーハの欠陥マップが生成された後、各欠陥場所及び相当する公知な欠陥のない参照場所が走査電子顕微鏡（SEM）により映され、位置及び地形データを集めると共に蓄積する。その後、このデータは境界分析および地形技術を使用してコア分級機により分析され、多数の所定の欠陥のコア分類の1つであるとして欠陥を分類する（コアは多数の一様なコア分類の1つ、又はユーザ

により望まれるコア分級機へプログラムされた別の分類であってもよい）。その後、欠陥は欠陥が配置されるコア分類と関連した特定の適応可能な分級機を使用して、可変の下位分類（すなわち、コア分級機により識別することができない下位分類）の任意数の1つにさらに分類され、効率のため、限定された規模で伝統的な分級機方法学を使用する。コア分級機又は特定の適用可能な分級機が欠陥をコア分類に分類することができない場合には、欠陥は伝統的な完全な分級機により分析される。

【0020】現在の方法学の特定の適応可能な分級機は、それがユーザにより1セットのサンプル欠陥画像で訓練され、欠陥属性に基づいて欠陥を分類するという点において、伝統的な分級機と同じ原則に基づいて動作する。しかし、多数の分類（例えば、10分類）の1つであるとしてすべての欠陥を分類しなければならない伝統的な完全な分級機と異なり、特定の適応可能な分級機はコア分類又はそれが関連する分類（例えば、1、2又は3分類）の欠陥だけを分析する。したがって、特定の適応可能な分級機は訓練されず、欠陥が多数の分類の1つであってもよいと考え、むしろ、特定の適応可能な分級機は訓練され、欠陥が特定のコア分類にあることを知る欠陥のその分析を開始する。例えば、表面の粒子、図1のコア分類3Eと関連した特定の適応可能な分級機は、欠陥が（3B～3D等の）パターン欠陥又は傷（3A）よりむしろ粒子であることを最初に決定する必要がないので、ユーザにより要求されるように表面の粒子の属性を迅速且つ確実に決定することができる。換言すれば、特定の適応可能な分級機は、全体の問題（10以上の1つとすることができるであろう欠陥）より解決し易い下位の問題（例えば、それが知る欠陥は2又は3の分類の1つである）で与えられる。特定の適応可能な分級機はあまり属性を考える必要がなく、ほとんどの状態でその判断に達するので、特定の適応可能な分級機の範囲のこの限界は伝統的な分級機より確実且つ迅速に欠陥を分類させる。すなわち、特定の適応可能な分級機は欠陥がどの分類にあるかを知っているため、それは最初、特定のセットの属性と考え、それが関連されるコア分類に不適切な属性と考えないであろう。

【0021】本発明は3つの異なるタイプの分級機の最良の属性を利用し、従来技術の方法学より迅速且つ確実に欠陥分類を実行する。それは所定数のコア分類への欠陥の早い最初の分類のためのコア分級機を使用し、コア分級機は欠陥の例を必要としないので、製造ラインの開始及び取り決めの間、動作することができる。それはまたコア分類内の部分的な改善のためコア分類と並行して特定の適応可能な分級機を使用する。本発明はさらに完全な分級機を使用するが、コア分類計画にそれを適合しない特定のタイプの欠陥のため限定された基準においてのみである。したがって、製作処理が完成し、関心の欠陥のタイプがもっと異なってくると、特定の適応可能な

完全な分級機はより詳細な欠陥分類を実行するために必要とされるように訓練されることができる。

【0022】図2は本発明の方法学により行われた自動的欠陥分類の概念フローチャートである。欠陥1は上述したように、パターン欠陥2A又は粒子欠陥2Bとして、コア分級機により広く分類され、7つの例示の一般的な欠陥のコア分類3A～3Gの1つにさらに配置される。コア分級機はコアの下位分類（図示せず）に欠陥をさらに配置することができ、コア分級機に予めプログラムされたモジュールを付加することによりユーザによって望まれているように、隣接する配線パターン、小粒子、大粒子、破線、狭いパターン等の間にブリッジング（すなわち、短絡）を含んでいる。その後、コア分類3A～3Gのそれぞれの欠陥はユーザにより所望のようにユーザによって訓練された特定の適応可能な分級機により任意数の可変な下位分類SC 4A1～4G2に分類される。代わりに、欠陥は、図2の断続線によって示されたような粒子又はパターン欠陥として広く分類された直後、特定の適応可能な分級機により4H等の下位分類に分類されることができる。特定の適応可能な分級機の前欠陥をコア分級機がどの程度分類するかについての判断は、最良の分類結果を得るための努力でのユーザの経験に基づいている。コア分級機又は特定の適応可能な分級機により分類することができない欠陥は適当な分類5へそれらを分類するユーザにより訓練された完全な分級機に送られる。

【0023】本発明の実施例は図3～6に示されている。図3に示されているように、欠陥のため検査される半導体ウェーハWは複数のパターン化された集積回路のダイ1000を有している。最初、欠陥マップは高速検査ツールでウェーハ表面を走査することによる等（CCD、レーザー又はSEMがこの目的のために用いられてもよい）の伝統的な技術で作られ、その後、通常、アルゴリズム及び又はグレースケール分析を含む統計的方法を使用して、欠陥を有する高い可能性を有するウェーハの疑わしい位置を確認する。

【0024】次に、図4A～4Cに示されているように、再検出が各疑わしい欠陥位置で行われ、欠陥の正確な位置を決定する。伝統的なCCDスキャナ又はSEMは疑わしい検出位置でパターン10を映すために使用されてもよく、それはその後、欠陥を有すると疑われない同じウェーハの近接又は他のダイの相当する位置で基準パターン20と比較される。疑わしい欠陥パターン10と基準パターン20の間で差異30が見出された場合には、疑わしい欠陥パターン10は結果であると決定され、創意に富む分析および分類が開始される。

【0025】本発明は図5に示されているように再調査ツール500で実行され、画像装置510を備え、それは好ましくは、カリフォルニア州のサンタクララ市のApplied Materials社から入手可能なSEMVision（登録商

標）又は共同出願中の出願に説明されているような高解像度SEMである。再調査ツール500は、好ましくは、ここに開示され、電子的に分析を行なうプロセッサ520と、プロセッサ520の分析の結果を表示するためのモニタとをさらに備えている。プロセッサ520は、伝統的なウェーハの基準設計データベース550と、半導体メモリ等のメモリデバイス540と通信することができる。

【0026】図6は創意に富む方法学の実施例のフローチャートであり、図7はプロセッサ520の機能的ブロック図である。段階600では、欠陥は上述したように画像装置510により映されている。次に、欠陥画像は、画像の境界分析及び又は画像の地形分析を実行することにより、共同出願中の出願に開示されているようにコア分級機により段階610で分析されている。段階620で、コア分級機700が欠陥を所定のコア分類（例えば、図2の分類2A～B又は3A～3G）に分類することができないと決定された場合には、欠陥画像は段階630で完全な分級機710により分析される。

【0027】コア分級機700が欠陥を分類する場合には、段階640で、コア分類の下位分類へのさらなる分類が要求されるかどうか決定される。そうでない場合には、分類手順は終了する。下位分類へのさらなる分類が必要である場合には、欠陥画像はコア分類2A、2B又は3A～3Gと関連した特定の適応可能な分級機720a～720eの1つによってさらに分析され、欠陥はコア分級機700により分類される（段階650参照）。例えば、コア分級機700が欠陥を表面の粒子（すなわち、コア分類3F）の粒子又は表面に埋め込まれた粒子（コア分類3E）として分類され、ユーザが粒子の寸法と共に形状を知りたい場合には、特定の適応可能な分級機720a～eの1つはユーザにより訓練され、それが公知の粒子欠陥を与えられ場合に粒子形状／寸法を決定し、その後、欠陥画像をさらに分析する。適当な特定の適応可能な分級機720a～720eが所望のように欠陥をうまく分類した場合には、分類手順が段階660で終了する。そうでなければ、欠陥は段階630で完全な分級機710によりさらに分析される。

【0028】減少された検査時間および現在の方法学の増加した分析信頼性の利点を達成するため、特定の適応可能な分級機720a～eはコア分類2A、2B、3A～3G、例えば、2又は3つのコア分類の限定数にそれぞれ関連している（すなわち、それから公知な欠陥を分析する）。各特定の適応可能な分級機720a～eは、下位分類の公知な欠陥の多数の例の属性（すなわち、特徴記述子）を特定の適応可能な分級機に供給することにより、下位分類の欠陥を認識するように訓練されている。作動では、特定の適応可能な分級機は欠陥の同じ属性を分析し、欠陥が下位分類にあるかどうかを決定する

ように分類される。

【0029】特定の適応可能な分級機720a～eは少数のコア分類の欠陥を分類するだけであるので、それぞれはそれらの分類からの欠陥だけ予想するように訓練されることができる。換言すれば、それぞれの特定の適応可能な分級機は欠陥画像の特定のセットの属性を見ることによりその分析を開始するように訓練されることができる。属性のこのリストは通常、伝統的な完全な分級機が分類されていない欠陥のその分析を開始する時に考えなければならない属性のリストより非常に短い。伝統的な完全な分級機は通常、40～100の属性を利用しているが、特定の適応可能な分級機は通常、約4～7の属性のみを考慮している。例えば、特定の適応可能な分級機は粒子欠陥であるとコア分級機700により決定された欠陥をさらに分類するために使用され、粒子の寸法、輝度、粗さ及び円磨度等の属性を最初に考えるように訓練されるであろう。対照的に、これらの属性は必ずしもパターン欠陥に関連しているとは限らないので、コア分級機700によりパターン欠陥であると決定された欠陥をさらに分類するための特定の適応可能な分級機は通常、輝度、円磨度及び粗さではなく、パターン寸法、他のパターンからの距離、パターン領域等のような属性を最初に考えるように訓練されるであろう。さらに、コア分級機700により傷であると決定された欠陥をさらに分類するための特定の適応可能な分級機は通常、散乱、伸長及び非円形等のさらに別のセットの属性を考えるように訓練されるであろう。したがって、比較的狭い範囲の欠陥を探すように訓練されることにより、特定の適応可能な分級機720a～eは、最小の検査時間でのコア分類の後の欠陥のさらなる分類を可能にする。

【0030】特定の適応可能な分級機720a～eは通常、判断ツリー技術又は多次元クラスタリング技術等の周知の伝統的な分級機技術を使用して訓練されるように構成されている。判断ツリー技術は通常、欠陥画像の最初の属性を分析する分級機を備え、その後、欠陥が分類されるまで、最初の属性の値を使用し、分析される次の属性を決定する等する。多次元クラスタリング技術では、分級機は公知なサンプル欠陥の多数の属性「x」を考えるように訓練され、各サンプルの欠陥はx次元空間にあるx座標を有する点として取扱われる。分級機はアルゴリズムを実行し、点をプロットし、それらが一緒にクラスタされているかどうかを決定する。その場合には、分級機は点を（例えば、長円体に）接続し、長円体内のすべての点はサンプル欠陥と同じ種類の欠陥を示すと考え。勿論、伝統的な分級機を作るために使用される技術は特定の適応可能な分級機を作るために使用可能である。

【0031】そのうえ、特定の適応可能な分級機は訓練されるよりむしろ欠陥を分類するために従う1セットのルールを与えられることができ、或いは、ルールベース

と学習した（すなわち、訓練された）分類技術の組合せを使用して動作可能である。例えば、コア分級機は欠陥を粒子として分類し、特定の適応可能な分級機が粒子の寸法を決定することになっている場合、特定の適応可能な分級機はプログラムされることができ、コア分級機により集められた情報を使用して粒子の直径を測定し、その後、アルゴリズムに従い、ユーザにより決定されるように、大粒子（例えば、 $x \mu m$ 以上の直径）、小粒子（ $x \mu m$ 以下の直径）等として粒子を分類する。代わりに、特定の適応可能な分級機は訓練され、特定の適応可能な分級機に多数の例の粒子欠陥を与えることにより粒子寸法を分析し、多次元クラスタリング技術を使用し、それらの寸法により粒子をプロットし、分級機により重要であると決定されるような小、中、大等の多数の寸法のカテゴリー（例えば、2、3又は4つの別のカテゴリー）にそれらをクラスタするように試みる。特定の適応可能な分級機が欠陥をクラスタできない場合、例えば、それらがクラスタされるよりむしろ均等又は任意に分配されることが分かった場合、その後、ユーザは上述したようなアルゴリズムを使用することにより、粒子を寸法で分類するための1セットのルールを設計することを必要とするであろう。

【0032】図8は図5に示された本発明の実施例を示すブロック図である。この実施例によれば、図5に示されるようなプロセッサ520はバス802又は情報を伝達する他の通信機構と、情報を処理するためバス802に結合された中央処理装置（CPU）804を備えている。プロセッサ520はまた、ランダムアクセスメモリ（RAM）又は他の動的記憶装置等の主メモリ806を備え、CPU804で実行される情報及び命令を記憶するためのバス802に結合される。主メモリ806はまた、仮の変数又はCPU804により実行される命令の実行の間の他の中間情報を記憶するために使用されてもよい。プロセッサ520はさらに読み出し専用メモリ（ROM）808又はバス802に結合される他の静的記憶装置を備え、CPU804のための静的情報及び命令を記憶する。磁気ディスク又は光学ディスク等の記憶装置810が供給され、バス802に結合され、情報及び命令を記憶する。記憶装置810はまた図5のメモリ540として有用であつてもよい。

【0033】プロセッサ520はバス802を介して陰極線管（CRT）等のモニター530に結合されてもよく、コンピュータユーザに情報を表示する。英数字及び他のキーを含む入力装置814はバス802に結合され、CPU804へ情報及び命令選択を通信する。ユーザ入力装置の別のタイプは、マウス、トラックボール、又はカーソル指示キー等のカーソル制御装置816であり、CPUへ指示情報及び命令選択を通信し、モニター330へのカーソル移動を制御する。

【0034】画像装置510（図3）は上述したよう

に、検査下の半導体ウェーハの画像を示すデータをバス 802 に入力する。そのようなデータは主メモリ 806 及び又は記憶装置 540 に記憶されてもよく、それが命令を実行する時に CPU 804 によって使用される。画像装置 510 はまたバス 802 を介して CPU 804 から命令を受け取る。

【0035】同様に、データベース（図 3）は上述したように、実質上欠陥のないウェーハを示すデータをバス 802 に入力する。そのようなデータは主メモリ 806 及び又は記憶装置 540 に記憶されてもよく、それが命令を実行する時に CPU 804 により使用される。

【0036】本発明はプロセッサ 520 の使用に関連し、半導体ウェーハの表面を欠陥のため検査する。本発明の 1 実施例によると、半導体ウェーハの表面の検査は表面欠陥の分類を含み、CPU 804 に応じてプロセッサ 520 により供給され、主メモリ 806 に含まれる 1 以上の命令の 1 以上のシーケンスを実行する。そのような命令は、記憶装置 810 等の別のコンピュータで読取り可能な媒体から主メモリに読まれてもよい。主メモリ 806 に含まれる命令のシーケンスの実行は上述した処理段階をプロセッサ 804 に行わせる。マルチ処理設備の 1 以上のプロセッサはまた使用され、主メモリ 806 に含まれた命令のシーケンスを実行する。代わりの実施例では、配線による回路がソフトウェア命令の代わり又はそれと共に使用され、本発明を実行することができる。したがって、本発明の実施例は配線回路とソフトウェアの特定の組合せに限定されない。装置のプログラミングは図 6 のフローチャートを供給された当業者であれば容易に行うことができる。

【0037】ここで使用されるコンピュータで読取り可能な媒体という用語は、実行のため CPU 804 に命令を供給することに関わる媒体に関連する。そのような媒体は、不揮発性媒体、揮発性媒体、及び伝送媒体を含むがそれらに限定されない多数の形をとってもよい。不揮発性の媒体は、例えば、記憶装置 810 等の光学又は磁気ディスクを含んでいる。揮発性媒体は主メモリ 806 等の動的メモリを含んでいる。伝送媒体は同軸ケーブル、銅線及びファイバーオプティクスを含み、バス 802 を構成する配線を含んでいる。伝送媒体はまた、無線周波数（RF）及び赤外線（IR）データ通信の間に発生されたもの等の音波又は光波の形をとることもできる。コンピュータで読取り可能な媒体の共通の形は、例えば、フロッピー（登録商標）ディスク、フレキシブルディスク、ハードディスク、磁気テープ、他の磁気媒体、CD-ROM、DVD、他の光学媒体、パンチカード、紙テープ、穴のパターンを有する他の理学媒体、ROM、PROM、及び EPROM、フラッシュ EPROM、他のメモリチップ又はカートリッジ、又はコンピュータが読取り可能な他の媒体を含んでいる。

【0038】コンピュータで読取り可能な媒体の各種形

式は、実行のため CPU 804 に 1 以上の命令の 1 以上のシーケンスを実行することを含んでいてもよい。例えば、最初、命令はリモートコンピュータの磁気ディスクに伝達されてもよい。リモートコンピュータはその動的メモリに命令をロードし、モデムを使用して電話線上で命令を送ることができる。プロセッサ 520 へのモデムは電話線でデータを受け取り、赤外線送信機を使用し、データを赤外線信号に変換することができる。バス 802 に結合された赤外線検出器は赤外線信号で実行されたデータを受け取り、バス 802 にデータを配置することができる。バス 802 はデータを主メモリ 806 に運び、CPU 804 は命令を検索すると共に実行する。主メモリ 806 により受け取られた命令は CPU 804 による実行前又は後のいずれかに記憶装置 810 に任意に記憶されてもよい。

【0039】創意に富んだ半導体ウェーハ検査技術は欠陥を粒子又はパターン欠陥等のコア分類に、そしてさらにコア分類の下位分類に迅速且つ確実に分類させる。創意に富んだ方法学は、ウェーハ表面を映すと共に境界分析及び又はその外形の地形測定を行なうことにより、欠陥をコア分類に分類するので、それは作動される前に各分類のための欠陥画像の例を要求しない。したがって、伝統的な分級機自体を使用する従来技術の欠陥分類とは異なり、本発明は製造ラインの開始及び取り決めの間に使用することができる。その上さらに、本発明は、少数のコア分類だけから欠陥を予想するように合わせて作られた特定の適応可能な分級機を使用し、ユーザが所望のように、欠陥を可変の下位分類に分類し、減少した検査時間及び増加した正確さを有する伝統的な分級機技術を使用する。したがって、本発明はコア分級機を使用することにより、伝統的なタイプの適応分級機の柔軟性と共に製造ラインの開始及び取り決めによる欠陥の分類を供給し、製造ラインが完成した時に欠陥の変更した下位分類の分類を可能にする。

【0040】本発明は半導体ウェーハの検査に適応可能であり、ミクロン以下の設計特徴を有する高密度の半導体デバイスの製造の間に半導体ウェーハの製造過程の検査に特に有用である。

【0041】本発明は、伝統的な材料、方法学及び機器を使用することにより実施可能である。したがって、そのような材料、機器及び方法学の詳細は詳細にここに示されていない。前の説明では、本発明の完全な理解を供給するため、特定の材料、構造、化学薬品、処理等、非常に多くの特定の詳細が示されている。しかし、当業者であれば分かるので、本発明は特に示された詳細に頼ることなく実施可能である。他の例では、本発明を不必要に不明確にしないために、周知の処理構造は詳細に説明されていない。

【0042】本発明の好適な実施例だけでなくその他用途の数例が本明細に示され、説明されている。本発明は

各種他の組合せ及び環境での使用が可能であり、ここに説明された創意に富んだ概念の範囲内の変更及び修正が可能であることは理解されるべきである。

【図面の簡単な説明】

【図 1】 共同出願中の発明による欠陥分類の概念フローチャートである。

【図 2】 本発明による欠陥分類の概念フローチャートである。

【図 3】 本発明を使用して検査される半導体ウェーハを示している。

【図 4 A】 本発明により検査される欠陥画像を表示したものである。

【図 4 B】 本発明により検査される欠陥画像を表示したものである。

【図 4 C】 本発明により検査される欠陥画像を表示したものである。

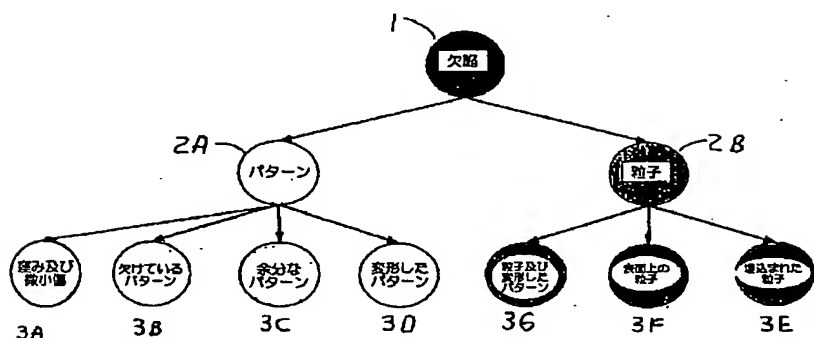
【図 5】 本発明による再調査ツールのブロック図である。

【図 6】 本発明による方法の一連の段階を示すフローチャートである。

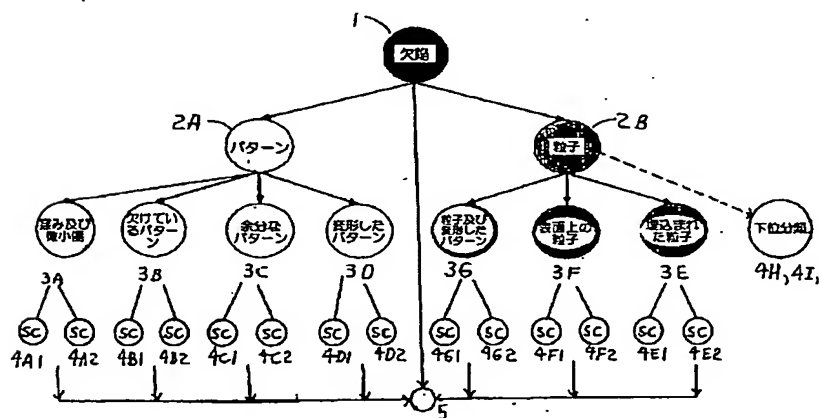
【図 7】 本発明による検査ツールの機能ブロック図である。

【図 8】 本発明の実施例を示すブロック図である。

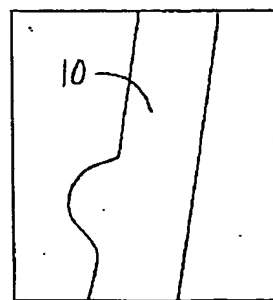
【図 1】



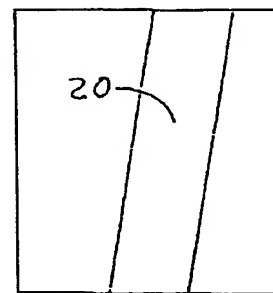
【図 2】



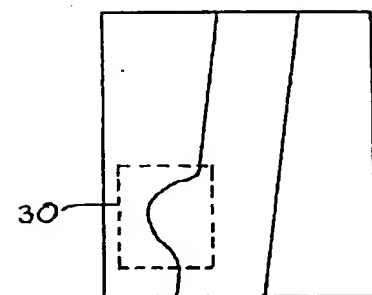
【図 4 A】



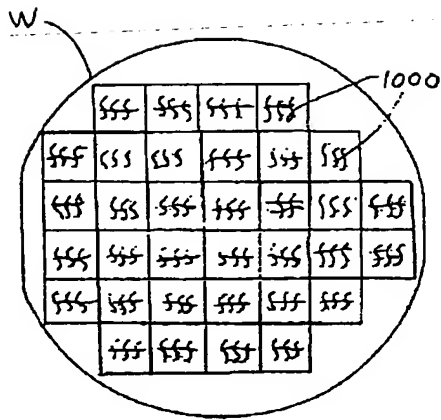
【図 4 B】



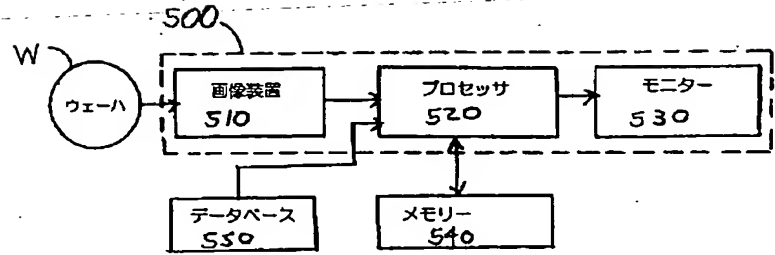
【図 4 C】



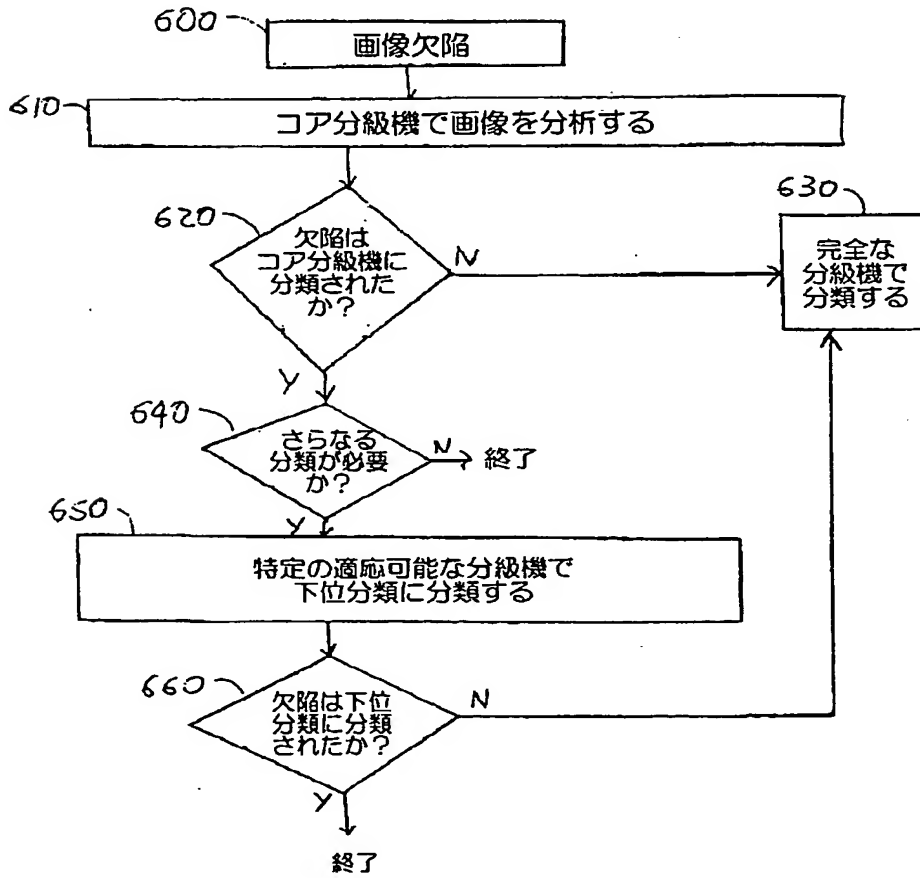
【図3】



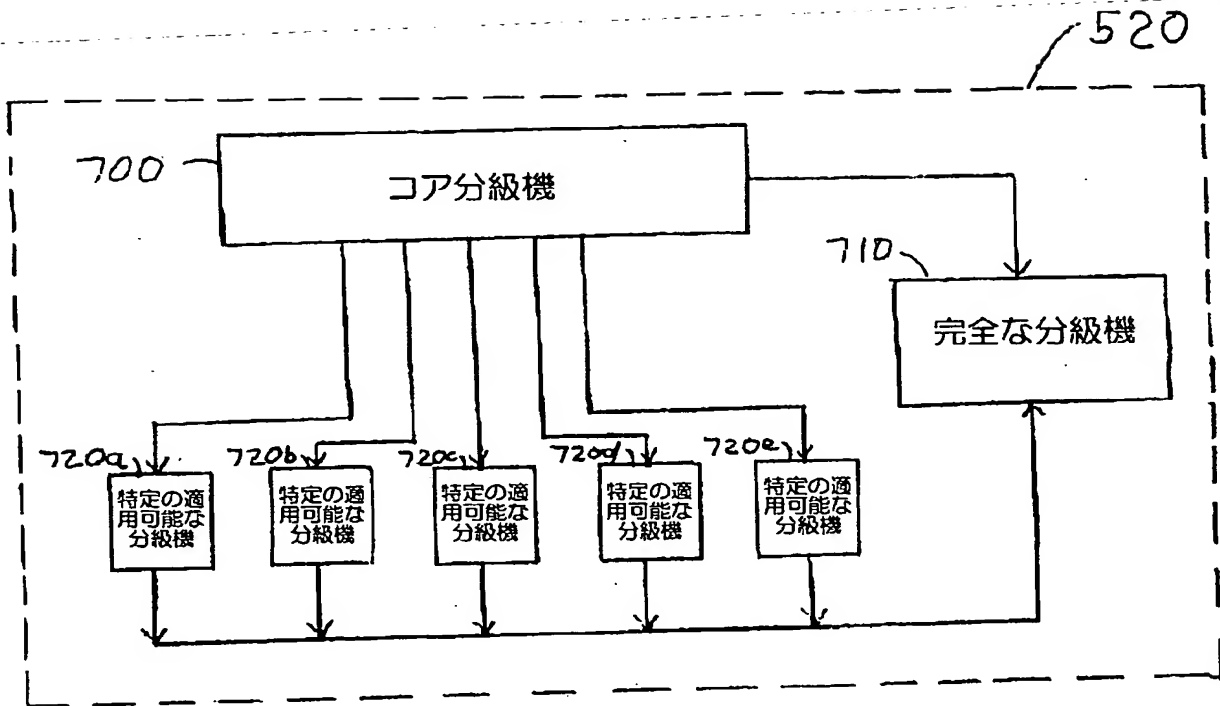
【図5】



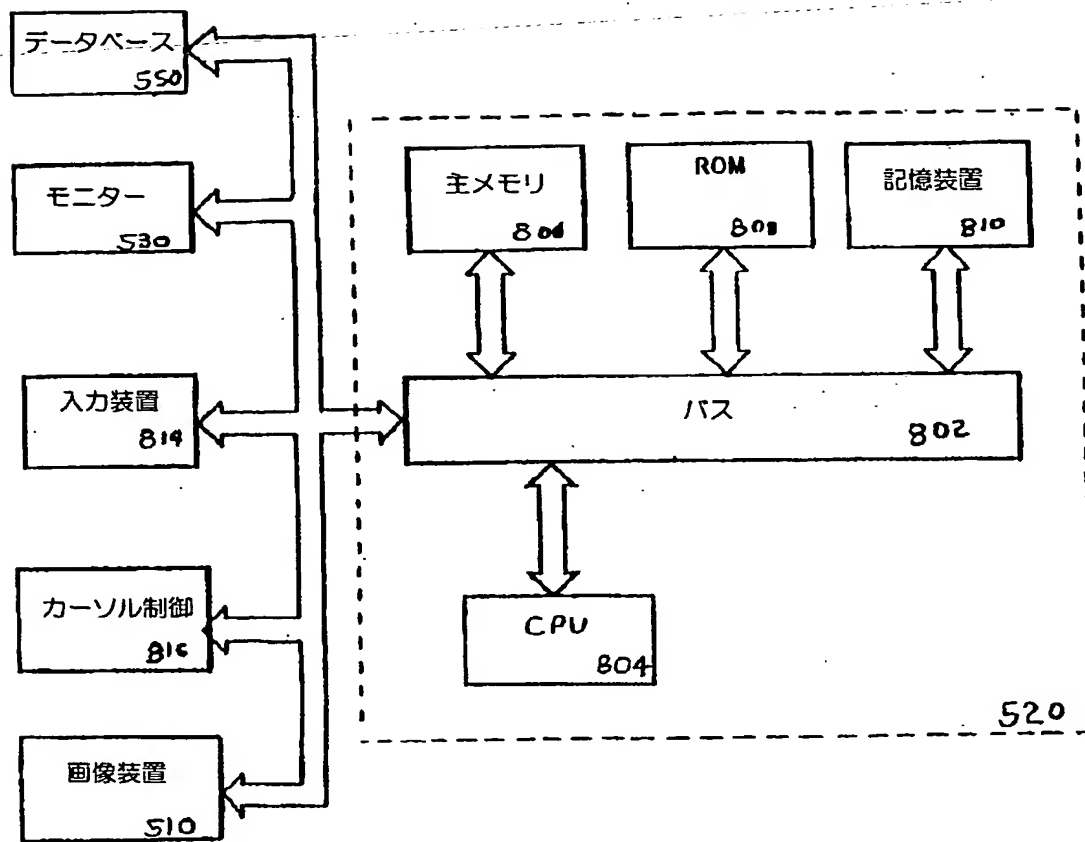
【図6】



【図7】



【図8】



HYBRID INVARIANT ADAPTIVE AUTOMATIC DEFECT CLASSIFICATION

FIELD OF THE INVENTION

The present invention relates to a method and apparatus for automatically classifying defects on the surface of an article. The invention has particular applicability for in-line inspection of semiconductor wafers during manufacture of high density semiconductor devices with submicron design features.

BACKGROUND ART

Current demands for high density and performance associated with ultra large scale integration require submicron features, increased transistor and circuit speeds and improved reliability. Such demands require formation of device features with high precision and uniformity, which in turn necessitates careful process monitoring, including frequent and detailed inspections of the devices while they are still in the form of semiconductor wafers.

Conventional in-process monitoring techniques employ an "inspection and review" procedure wherein the surface of the wafer is initially scanned by a high-speed, relatively low-resolution inspection tool; for example, an opto-electric converter such as a CCD (charge-coupled device) or a laser. Statistical methods are then employed to produce a defect map showing suspected locations on the wafer having a high probability of a defect. If the number and/or density of the potential defects reaches a predetermined level, an alarm is sounded, indicating that a more detailed look at the potential defect sites is warranted. This technique is known as "total density monitoring" of defects and produces a statistic called the "total defect density".

When the defect density reaches a predetermined level, a review of the affected wafers is warranted. After a redetection procedure is carried out, using the defect map, to positively determine the presence of defects, a more detailed review procedure is carried out on the individual defect sites, such as scanning with a CCD to produce a relatively high-resolution image. The defect image is then analyzed to determine the nature of the defect (e.g., a defective pattern, a particle, or a scratch).

Since it has recently been recognized that monitoring "classified defect density", i.e., the number of defects of several different types, or "classes" of defects, is preferable to monitoring total defect density, various methods for classification of defects have been introduced. Most of these conventional methods, called "classic classifiers" herein, employ pattern recognition techniques wherein a set of sample defects is acquired, imaged and analyzed for particular characteristics or "predicates" (e.g., brightness, roughness, size, color), which predicates are fed into a "black box" (e.g., a neural net) and used to train the classifier to recognize different types of defects by the defects' predicates.

Disadvantageously, the efficiency of these methods is reduced because there is no agreed-upon set of defect classes. Different semiconductor fabricators consider different defects to be important and, therefore, use different sets of defect classes tailored to their specific needs. Thus, they require many examples of defect images to be obtained for each defect class prior to becoming operational. Consequently, typical prior art systems cannot be used during start-up and ramp-up of a production line. Furthermore, because such classifiers, also referred to as "full classifiers" herein, need to discriminate between all defect types required to be classified (e.g., 10 or more classes of defects), a large number of predicates must be considered when classifying any defect, thus increasing inspection time and reducing production throughput.

To address the above-mentioned problems associated with full classifiers, an invariant core classifier ("core classifier" herein) has recently been introduced in the defect review system marketed as the SEMVisionTM, available from Applied Materials of Santa Clara, California. Such a core classifier is described in copending U.S. Patent Application Serial No. 09/111,454, filed July 8, 1998, entitled "Automatic Defect Classification With Invariant Core Classes", the entire disclosure of which is hereby incorporated by reference.

According to the methodology of the copending application, after a defect map of a semiconductor wafer has been generated, each defect site and a corresponding known non-defective reference site is imaged by a scanning electron microscope (SEM) to gather and store location and topographical data. The image is then analyzed, as by performing boundary analyses and/or topographical measurements, to classify the defect as being in one of a number (e.g., seven) of invariant core classes of defect, and further classified as being in one of an arbitrary number of core sub-classes as desired by the user by adding pre-programmed modules onto the core classifier.

Fig. 1 is a conceptual flow chart of automatic defect classification into core classes performed by the methodology of the copending application. A defect 1 is classified broadly as a pattern defect 2A or a particle defect 2B, and further placed into one of seven exemplary invariant core classes of defects: craters and microscratches on the wafer surface 3A, a missing pattern on the surface 3B, an extra pattern on the surface 3C, a deformed pattern on the surface 3D, a particle on the surface 3E, a particle embedded in the surface 3F, or a particle and a deformed pattern on the surface 3G. Arbitrary core subclasses may include bridging (i.e., short circuiting) between neighboring wiring patterns, a small particle, a large particle, a broken line, a narrow pattern, etc.

The invariant core classification technique of the copending application enables defects to be separately and reliably classified as particle or pattern defects, and as on-surface or below-surface (embedded) defects. It also provides early quantification and notification of these meaningfully classified defects, thereby facilitating investigation of the causes of the defects, and enabling early corrective action to be implemented.

The core classifier of the copending application is a "rule-based" classifier in that it classifies defects by collecting defect information (i.e., imaging the wafer surface and performing boundary analysis and/or topographical measurement of its features) then following a set of rules programmed *a priori* (i.e., beforehand). Thus, it does not need to be trained, as do classic classifiers, and so does not require examples of defect images for each class prior to being operational. Therefore, unlike prior art defect classification systems, the core classifier of the copending application can be used during the start-up and ramp-up of a production line.

While core classifiers as described in the copending application address many of the shortcomings of conventional classic classifiers, core classifiers may not be suitable for separating defects into every class deemed important by a user since, as rule-based classifiers, they cannot be easily adapted to recognize new classes of defects. Specifically, the user may require refinements within the invariant core classes (since different process lines may be sensitive to different defects from one to another) other than the core subclasses discussed above available as pre-programmed modules to be added to the core classifier. Furthermore, the user may require refinements that cannot be discerned by the core classifier. For example, if the core classifier classifies a defect as a particle on the surface (core class 3F in Fig. 1), and the user wishes to know the shape of the particle in combination with its size, another technique must be used to obtain this size information,

which is helpful in pinpointing the source of the particle, since different processes tend to produce different particle shapes and sizes. Additionally, "exotic" defects that do not fall into any of the core classes cannot be classified by a core classifier. For example, if a process is introduced that results in a new type of defect, the existing core classes will be irrelevant in relation to the new defect, and the core classifier will not be able to classify the new defect unless the new defect is added as a core class.

There exists a need to quickly and meaningfully review semiconductor wafers and automatically classify defects using a core classifier, then further classify the defects into subclasses within a core class desired by the user in order to identify processes causing defects, thereby enabling early corrective action to be taken. This need is becoming more critical as the density of surface features, die sizes, and number of layers in devices increase, requiring the number of defects to be drastically reduced to attain an acceptable manufacturing yield.

SUMMARY OF THE INVENTION

An object of the present invention is to provide a method and apparatus for automatic, fast, reliable and complete classification of defects in semiconductor wafers.

According to the present invention, the foregoing and other objects are achieved in part by a method of classifying a defect on the surface of an article, which method comprises imaging the surface, classifying the defect as being in one of a predetermined number of core classes of defects using a core classifier, and classifying the defect as being in one of an arbitrary number of variant subclasses using a specific adaptive classifier associated with the one core class. If the core classifier cannot classify the defect into a core class, a full classifier is used to classify the defect.

A further aspect of the present invention is an apparatus for carrying out the steps of the above method.

A still further aspect of the present invention is a computer-readable medium bearing instructions for automatically classifying a defect on the surface of an article, the instructions, when executed, being arranged to cause one or more processors to perform the steps of the above method.

Additional objects and advantages of the present invention will become readily apparent to those skilled in this art from the following detailed description, wherein only the preferred embodiment of the invention is shown and described, simply by way of

illustration of the best mode contemplated for carrying out the invention. As will be realized, the invention is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the invention. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

BRIEF DESCRIPTION OF THE DRAWINGS

Reference is made to the attached drawings, wherein elements having the same reference numeral designations represent like elements throughout, and wherein:

Figure 1 is a conceptual flow chart of defect classification according to the copending invention.

Figure 2 is a conceptual flow chart of defect classification according to the present invention.

Figure 3 illustrates a semiconductor wafer to be inspected using the present invention.

Figures 4A-4C are representations of images of a defect to be inspected by the present invention.

Figure 5 is a block diagram of a review tool according to the present invention.

Figure 6 is a flow chart illustrating sequential steps of a method according to the present invention.

Figure 7 is a function block diagram of an inspection tool according to the present invention.

Figure 8 is a block diagram that illustrates an embodiment of the invention.

DESCRIPTION OF THE INVENTION

Core classifiers, such as described in copending U.S. Application 09/111,454, enable defects to be quickly and reliably classified as particle or pattern defects, and as on-surface or below-surface (embedded) defects, and are operable during start-up and ramp-up of a production line. However, core classifiers cannot classify certain types of defects within its core classes, and are not easily modified by the user to add defect classes. Conventional classic classifiers, on the other hand, are fully adaptable to classify any type of defect as trained by the user, but require many examples of defect images to be obtained for each defect class prior to becoming operational. Thus, classic classifiers cannot be

used during start-up and ramp-up of a production line. Furthermore, classic-style full classifiers are slower than core classifiers, and generally less accurate. The present invention addresses and solves these problems by providing a hybrid defect classification methodology utilizing a core classifier together with limited-range classic-style classifiers for fast and reliable classification of defects into invariant core classes and variant subclasses as needed by the user. Thus, the present methodology enables ready identification of processes causing defects, thereby enabling early corrective action to be taken.

According to certain embodiments of the methodology of the present invention, after a defect map of a semiconductor wafer has been generated, each defect site and a corresponding known non-defective reference site is imaged by a scanning electron microscope (SEM) to gather and store location and topographical data. This data is then analyzed by a core classifier using boundary analysis and topographical techniques to classify the defect as being in one of a number of predetermined core classes of defect (the core class may be one of a number of invariant core classes or a variant class programmed into the core classifier as desired by the user). The defect is then further classified into one of an arbitrary number of variable subclasses (i.e., subclasses that cannot be discerned by the core classifier) using a "specific adaptive classifier" associated with the core class in which the defect has been placed, which employs classic classifier methodology on a limited scale, for efficiency. If the core classifier or specific adaptive classifier cannot classify the defect into a core class, the defect is analyzed by a conventional full classifier.

The specific adaptive classifier of the present methodology operates based on the same principles as a classic classifier, in that it is trained by the user with a set of sample defect images to classify a defect based on the defect's predicates. However, unlike conventional full classifiers that must classify every defect as being in one of many classes (e.g., 10 classes), the specific adaptive classifier analyzes only defects of the core class or classes with which it is associated (e.g., 1, 2 or 3 classes). Thus, the specific adaptive classifier is not trained to consider that the defect may be in one of many classes; rather, the specific adaptive classifier is trained to start its analysis of the defect "knowing" that the defect is in a particular core class. For example, a specific adaptive classifier associated with the "particle on surface" core class 3E in Fig. 1 is able to quickly and accurately determine the attributes of a particle on the surface as required by the user, because it does not need to first determine that the defect is a particle rather than a pattern

defect (such as 3B-3D) or a scratch (3A). In other words, the specific adaptive classifier is presented with a "sub-problem" (e.g., a defect that it knows is in one of, say, 2 or 3 classes) that is easier to solve than the total problem (a defect that could be in one of 10 or more classes). This limit on the domain of the specific adaptive classifier enables it to classify defects more reliably and faster than a conventional classic classifier, since the specific adaptive classifier needs to consider less predicates to reach its decision in most situations. That is, because the specific adaptive classifier knows what class the defect is in, it will consider a particular set of predicates first and not consider predicates irrelevant to the core class with which it is associated.

The present invention utilizes the best attributes of three different types of classifiers to perform defect classification more quickly and reliably than prior art methodologies. It employs a core classifier for fast initial classification of defects into a predetermined number of core classes, the core classifier being able to work during start-up and ramp-up of a production line because it does not require examples of defects. It also uses specific adaptive classifiers in parallel with the core classes for local refinement within a core class. The present invention further employs a full classifier, but only on a limited basis for special types of defects that do not fit in to the core classification scheme. Thus, as the fabrication process matures and the types of defects of interest become more diverse, the specific adaptive and full classifiers can be trained as needed to perform more detailed defect classification.

Fig. 2 is a conceptual flow chart of automatic defect classification performed by the methodology of the present invention. A defect 1 is classified broadly by a core classifier, as discussed above, as a pattern defect 2A or a particle defect 2B, and further placed into one of seven exemplary invariant core classes of defects 3A-3G. The core classifier can further place the defects in core subclasses (not shown) including bridging (i.e., short circuiting) between neighboring wiring patterns, a small particle, a large particle, a broken line, a narrow pattern, etc. as desired by the user by adding pre-programmed "modules" onto the core classifier. Defects in each of the core classes 3A-3G are then classified into an arbitrary number of variable subclasses SC 4A1-4G2 as desired by the user by specific adaptive classifiers trained by the user. Alternatively, defects can be classified by a specific adaptive classifier into a subclass such as 4H directly after being broadly classified as a particle or pattern defect, as shown by the dashed line in Fig. 2. The decision as to how far the core classifier classifies a defect

before the specific adaptive classifiers are used is based on the experience of the user in an effort to obtain the best classification results. Defects that cannot be classified by the core classifier or the specific adaptive classifiers are sent to a full classifier trained by the user that classifies them into an appropriate class 5.

An embodiment of the present invention is illustrated in Figs. 3-6. As shown in Fig. 3, a semiconductor wafer W to be inspected for defects has a plurality of patterned integrated circuit dies 1000. Initially, a defect map is produced by conventional techniques, such as by scanning the surface of a wafer with a high-speed inspection tool (a CCD, a laser or an SEM may be employed for this purpose), then using statistical methods, typically involving algorithms and/or grey-scale analysis, to identify suspected locations on the wafer having a high probability of having a defect.

Next, as shown in Figs. 4A-4C, a redetection procedure is carried out at each suspected defect location to determine the exact location of the defect. A conventional CCD scanner or an SEM may be used to image a pattern 10 at a suspected defect location, which is then compared to a reference pattern 20 at a corresponding location on an adjacent or other die on the same wafer which is not suspected of having a defect. If a difference 30 is found between the suspected defective pattern 10 and the reference pattern 20, the suspected defective pattern 10 is determined to be a defect, and the inventive analysis and classification commences.

The present invention is implemented at a review tool 500, as shown in Fig. 5, comprising an imager 510, preferably a high resolution SEM such as the SEMVision™ available from Applied Materials, Santa Clara, California, or as described in the copending application. Review tool 500 further comprises a processor 520, which preferably performs the analysis disclosed herein electronically, and a monitor 530 for displaying results of the analyses of processor 520. Processor 520 can be in communication with a conventional wafer reference design database 550 and a memory device 540, such as a semiconductor memory.

Fig. 6 is a flow chart of an embodiment of the inventive methodology, and Fig. 7 is a functional block diagram of processor 520. At step 600, the defect is imaged by imager 510 as discussed above. Next, the defect image is analyzed in step 610 by a core classifier 700 as disclosed in the copending application, by performing a boundary analysis of the image and/or a topographical analysis of the image. If it is determined at step 620 that

core classifier 700 cannot classify the defect into a predetermined core class (e.g., classes 2A-B or 3A-3G in Fig. 2), the defect image is analyzed by full classifier 710 at step 630.

If core classifier 700 classifies the defect, it is determined at step 640 whether further classification into a subclass of a core class is required. If not, the classification procedure ends. If further classification into a subclass is needed, the defect image is further analyzed by one of specific adaptive classifiers 720a-720e associated with the core class 2A, 2B or 3A-3G into which the defect was classified by core classifier 700 (see step 650). For example, if core classifier 700 classifies the defect as a particle on the surface (i.e., core class 3F) or a particle embedded in the surface (core class 3E) and the user wishes to know the shape in combination with the size of the particle, one of the specific adaptive classifiers 720a-e, that has been trained by the user to determine particle shape/size if it is given a known particle defect, then further analyzes the defect image. If the appropriate specific adaptive classifier 720a-720e successfully classifies the defect as desired, the classification procedure ends at step 660. If not, the defect is further analyzed by full classifier 710 at step 630.

To achieve the advantages of reduced inspection time and increased classification reliability of the present methodology, specific adaptive classifiers 720a-e are each associated with (i.e., analyze known defects from) a limited number of core classes 2A, 2B, 3A-3G, e.g., two or three core classes. Each specific adaptive classifier 720a-e is trained to identify defects of a subclass by providing the specific adaptive classifier a set of predicates (i.e., feature descriptors) of a number of examples of known defects of the subclass. In operation, the specific adaptive classifier analyzes the same predicates of a defect to be classified to determine if the defect falls in the subclass.

Because specific adaptive classifiers 720a-e only classify defects of a small number of core classes, each can be trained to expect only defects from those classes; in other words, each specific adaptive classifier can be trained to begin its analysis by looking at a particular set of predicates of the defect image. This list of predicates is generally much shorter than the list of predicates a conventional full classifier must consider as it begins its analysis of an unclassified defect. Specific adaptive classifiers typically consider only about 4-7 predicates, while conventional full classifiers typically utilize 40-100 predicates. For example, a specific adaptive classifier used to further classify a defect determined by core classifier 700 to be a particle defect would have been trained to first consider predicates such as size, brightness, roughness and roundness of the

particle. In contrast, a specific adaptive classifier for further classifying a defect determined by core classifier 700 to be a pattern defect would typically be trained to first consider predicates such as pattern size, distance from other patterns, area of the pattern, etc.; but not brightness, roundness and roughness, since these predicates are not necessarily relevant to pattern defects. Additionally, a specific adaptive classifier for further classifying defects determined to be scratches by core classifier 700 would typically be trained to consider yet another different set of predicates, such as scatter, elongation and non-circularity. Thus, by being trained to look for a relatively narrow range of defects, specific adaptive classifiers 720a-e enable further classification of defects after core classification with a minimum of inspection time.

Specific adaptive classifiers 720a-e are typically configured to be trained using well-known classic classifier techniques, such as a "decision tree" technique or a "multidimensional clustering" technique. The decision tree technique typically comprises the classifier analyzing a first predicate of the defect image, then using the value of the first predicate to determine the next predicate to be analyzed, and so on until the defect is classified. In the multidimensional clustering technique, the classifier is trained to consider a number of predicates "x" of the known sample defects, each sample defect being treated as a point having x coordinates hanging in x-dimensional space. The classifier executes algorithms to plot the points and determine whether they are clustered together. If so, the classifier connects the points (e.g., with an ellipsoid) and deems that every point inside the ellipsoid represents the same kind of defect as the sample defects. Of course, any technique used to construct a classic classifier can be used to construct a specific adaptive classifier.

Moreover, a specific adaptive classifier can be given a set of rules to follow to classify defects, rather than being trained, or can operate using a combination of rule-based and learned (i.e., trained) classification techniques. For example, if the core classifier classifies a defect as a particle, and the specific adaptive classifier is to determine the size of the particle, the specific adaptive classifier can be programmed to "measure" the particle's diameter using information gathered by the core classifier and then follow an algorithm to classify the particle as a "large particle" (e.g., above $x\mu\text{m}$ diameter), "small particle" (e.g., below $x\mu\text{m}$ diameter), etc. as determined by the user. Alternatively, the specific adaptive classifier can be trained to analyze particle size by giving the specific adaptive classifier a number of examples of particle defects, using a multidimensional

clustering technique to plot the particles according to their size, and attempting to cluster them into a number of size categories (e.g., two, three, or four separate categories) such as "small", "medium", "large", etc. as determined to be significant by the classifier. If the specific adaptive classifier cannot cluster the defects; e.g., if they are found to be evenly or randomly distributed rather than clustered, then the user will need to design a set of rules for classifying the particles by size, as by using an algorithm as discussed above.

Figure 8 is a block diagram that illustrates an embodiment of the invention shown in Fig. 5. According to this embodiment, processor 520, as shown in Fig. 5, includes a bus 802 or other communication mechanism for communicating information, and a central processing unit (CPU) 804 coupled with bus 802 for processing information. Processor 520 also includes a main memory 806, such as a random access memory (RAM) or other dynamic storage device, coupled to bus 802 for storing information and instructions to be executed by CPU 804. Main memory 806 also may be used for storing temporary variables or other intermediate information during execution of instructions to be executed by CPU 804. Processor 520 further includes a read only memory (ROM) 808 or other static storage device coupled to bus 802 for storing static information and instructions for CPU 804. A storage device 810, such as a magnetic disk or optical disk, is provided and coupled to bus 802 for storing information and instructions. Storage device 810 may also serve as memory 540 in Fig. 5.

Processor 520 may be coupled via bus 802 to monitor 530, such as a cathode ray tube (CRT), for displaying information to a computer user. An input device 814, including alphanumeric and other keys, is coupled to bus 802 for communicating information and command selections to CPU 804. Another type of user input device is cursor control 816, such as a mouse, a trackball, or cursor direction keys for communicating direction information and command selections to CPU 804 and for controlling cursor movement on monitor 330.

Imager 510 (Fig. 3) inputs data representative of images of a semiconductor wafer under inspection, as discussed above, to bus 802. Such data may be stored in main memory 806 and/or storage device 540, and used by CPU 804 as it executes instructions. Imager 510 may also receive instructions via bus 802 from CPU 804.

Likewise, database 550 (Fig. 3) inputs data representative of a substantially defectless wafer, as discussed above, to bus 802. Such data may be stored in main memory 806 and/or storage device 540, and used by CPU 804 as it executes instructions.

The invention is related to the use of processor 520 for inspecting the surface of a semiconductor wafer for defects. According to one embodiment of the invention, inspection of the surface of a semiconductor wafer, including classification of surface defects, is provided by processor 520 in response to CPU 804 executing one or more sequences of one or more instructions contained in main memory 806. Such instructions may be read into main memory 806 from another computer-readable medium, such as storage device 810. Execution of the sequences of instructions contained in main memory 806 causes processor 804 to perform the process steps described above. One or more processors in a multi-processing arrangement may also be employed to execute the sequences of instructions contained in main memory 806. In alternative embodiments, hard-wired circuitry may be used in place of or in combination with software instructions to implement the invention. Thus, embodiments of the invention are not limited to any specific combination of hardware circuitry and software. The programming of the apparatus is readily accomplished by one of ordinary skill in the art provided with the flow chart of Fig. 6.

The term "computer-readable medium" as used herein refers to any medium that participates in providing instructions to CPU 804 for execution. Such a medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media include, for example, optical or magnetic disks, such as storage device 810. Volatile media include dynamic memory, such as main memory 806. Transmission media include coaxial cable, copper wire and fiber optics, including the wires that comprise bus 802. Transmission media can also take the form of acoustic or light waves, such as those generated during radio frequency (RF) and infrared (IR) data communications. Common forms of computer-readable media include, for example, a floppy disk, a flexible disk, hard disk, magnetic tape, any other magnetic medium, a CD-ROM, DVD, any other optical medium, punch cards, paper tape, any other physical medium with patterns of holes, a RAM, a PROM, and EPROM, a FLASH-EPROM, any other memory chip or cartridge, or any other medium from which a computer can read.

Various forms of computer-readable media may be involved in carrying out one or more sequences of one or more instructions to CPU 804 for execution. For example, the instructions may initially be borne on a magnetic disk of a remote computer. The remote computer can load the instructions into its dynamic memory and send the instructions over

a telephone line using a modem. A modem local to processor 520 can receive the data on the telephone line and use an infrared transmitter to convert the data to an infrared signal. An infrared detector coupled to bus 802 can receive the data carried in the infrared signal and place the data on bus 802. Bus 802 carries the data to main memory 806, from which CPU 804 retrieves and executes the instructions. The instructions received by main memory 806 may optionally be stored on storage device 810 either before or after execution by CPU 804.

The inventive semiconductor wafer inspection technique enables defects to be quickly and reliably classified into core classes such as particle or pattern defects, and further into subclasses of the core classes. Since the inventive methodology classifies defects into core classes by imaging the wafer surface and performing boundary analysis and/or topographical measurement of its features, it does not require examples of defect images for each class prior to being operational. Therefore, unlike prior art defect classification systems using classic classifiers alone, the present invention can be used during the start-up and ramp-up of a production line. Furthermore, the present invention utilizes specific adaptive classifiers, tailored to expect defects from only a small number of core classes, to classify defects into variable subclasses as desired by the user, employing classic classifier techniques with reduced inspection time and increased accuracy. Thus, the present invention provides classification of defects upon start-up and ramp-up of a production line by using a core classifier, along with the flexibility of classic-style adaptive classifiers to allow the classification of varied subclasses of defects as the production line matures.

The present invention is applicable to the inspection of any semiconductor wafer, and is especially useful for in-process inspection of semiconductor wafers during manufacture of high-density semiconductor devices with submicron design features.

The present invention can be practiced by employing conventional materials, methodology and equipment. Accordingly, the details of such materials, equipment and methodology are not set forth herein in detail. In the previous descriptions, numerous specific details are set forth, such as specific materials, structures, chemicals, processes, etc., in order to provide a thorough understanding of the present invention. However, as one having ordinary skill in the art would recognize, the present invention can be practiced without resorting to the details specifically set forth. In other instances, well known

processing structures have not been described in detail, in order not to unnecessarily obscure the present invention.

Only the preferred embodiment of the invention and but a few examples of its versatility are shown and described in the present disclosure. It is to be understood that the invention is capable of use in various other combinations and environments and is capable of changes or modifications within the scope of the inventive concept as expressed herein.

What is Claimed Is:

1. A method of classifying a defect on the surface of an article, which method comprises:

imaging the surface to form a defect image;

classifying the defect as being in one of a predetermined number of core classes of defects using a core classifier; and

classifying the defect as being in one of an arbitrary number of variant subclasses using a specific adaptive classifier associated with the one core class.

2. The method according to claim 1, comprising classifying the defect as being in one of an arbitrary number of variant classes using a full classifier when the core classifier cannot classify the defect into one of the core classes.

3. The method according to claim 1, wherein the core classes of defects comprise a pattern defect and a particle defect.

4. The method according to claim 3, wherein the core classes of defects comprise a missing pattern on the surface, an extra pattern on the surface, a particle on the surface, a particle embedded in the surface, and microscratches on the surface.

5. The method according to claim 4, wherein the variant subclasses of defects comprise a particle of a predetermined size on the surface or embedded in the surface.

6. The method according to claim 1, comprising providing a plurality of specific adaptive classifiers, each being associated with less than the predetermined number of core classes.

7. The method according to claim 1, comprising performing a boundary analysis of the defect image, a topographical analysis of the defect image, or both to classify the defect into the one core class.

8. The method according to claim 7, comprising:

training the specific adaptive classifier to identify defects of the variant subclass based on an exemplary set of known predicates of defects of the variant subclass; and

analyzing predicates associated with the defect image, using the specific adaptive classifier, to classify the defect into one of the variant subclasses.

9. The method according to claim 8, comprising training the specific adaptive classifier using decision tree or multidimensional clustering techniques.

10. An apparatus for classifying a defect on the surface of an article, comprising:

an imager to produce an image of the defect; and

a processor comprising:

a core classifier for classifying the defect as being in one of a predetermined number of core classes of defects, and

a specific adaptive classifier associated with the one core class for classifying the defect as being in one of an arbitrary number of variant subclasses.

11. The apparatus according to claim 10, wherein the processor further comprises a full classifier for classifying the defect as being in one of an arbitrary number of variant classes when the core classifier cannot classify the defect into one of the core classes.

12. The apparatus according to claim 10, wherein the core classes of defects comprise a pattern defect or a particle defect.

13. The apparatus according to claim 12, wherein the core classes of defects comprise a missing pattern on the surface, an extra pattern on the surface, a particle on the surface, a particle embedded in the surface, and microscratches on the surface.

14. The apparatus according to claim 13, wherein the variant subclasses of defects comprise a particle of a predetermined size on the surface or embedded in the surface.

15. The apparatus according to claim 10, comprising a plurality of specific adaptive classifiers, each being associated with less than the predetermined number of core classes.

16. The apparatus according to claim 10, wherein the processor is further configured to perform a boundary analysis of the defect image, a topographical analysis of the defect image, or both to classify the defect into the one core class.

17. The apparatus according to claim 16,
wherein the specific adaptive classifier is trainable to identify defects of the variant subclass based on an exemplary set of known predicates of defects of the variant subclass;
and

wherein the specific adaptive classifier is for analyzing predicates associated with the defect image to classify the defect into one of the variant subclasses.

18. The method according to claim 17, wherein the specific adaptive classifier is trainable using decision tree or multidimensional clustering techniques.

19. A specific adaptive classifier for classifying a defect on the surface of an article as being in one of an arbitrary number of variant subclasses of a core defect class, responsive to a core classifier classifying the defect as being in the core class.

20. The specific adaptive classifier of claim 19, wherein the specific adaptive classifier is trainable to identify defects of the variant subclass based on an exemplary set of known predicates of defects of the variant subclass; and

wherein the specific adaptive classifier is for analyzing predicates associated with an image of the defect to classify the defect into one of the variant subclasses.

21. The specific adaptive classifier of claim 20, wherein the specific adaptive classifier is trainable using decision tree or multidimensional clustering techniques.

22. A computer-readable medium bearing instructions for automatically classifying a defect on the surface of an article, said instructions, when executed, being arranged to cause one or more processors to perform the steps of:

imaging the surface to form a defect image;

classifying the defect as being in one of a predetermined number of core classes of defects; and

classifying the defect as being in one of an arbitrary number of variant subclasses based on the classification of the defect as being in the one core class.

23. The computer-readable medium according to claim 22, wherein the instructions, when executed, are arranged to cause the one or more processors to perform the step of classifying the defect as being in one of an arbitrary number of variant classes when the one or more processors cannot classify the defect into one of the core classes.

24. The computer-readable medium according to claim 22, wherein the core classes of defects comprise a pattern defect and a particle defect.

25. The computer-readable medium according to claim 24, wherein the core classes of defects comprise a missing pattern on the surface, an extra pattern on the surface, a particle on the surface, a particle embedded in the surface, and microscratches on the surface.

26. The computer-readable medium according to claim 25, wherein the variant subclasses of defects comprise a particle of a predetermined size on the surface or embedded in the surface.

27. The computer-readable medium according to claim 22, wherein the instructions, when executed, are arranged to cause the one or more processors to perform a boundary analysis of the defect image, a topographical analysis of the defect image, or both to classify the defect into the one core class.

28. The computer-readable medium according to claim 26, wherein the instructions, when executed, are arranged to cause the one or more processors to perform the steps of:

learning to identify defects of the variant subclass based on an exemplary set of known predicates of defects of the variant subclass; and

analyzing predicates associated with the defect image to classify the defect into one of the variant subclasses.

29. The computer-readable medium according to claim 28, wherein the instructions, when executed, are arranged to cause the one or more processors to perform the learning step using decision tree or multidimensional clustering techniques.

30. The computer-readable medium according to claim 22, wherein the instructions, when executed, are arranged to cause the one or more processors to perform the step of imaging the surface with a scanning electron microscope.

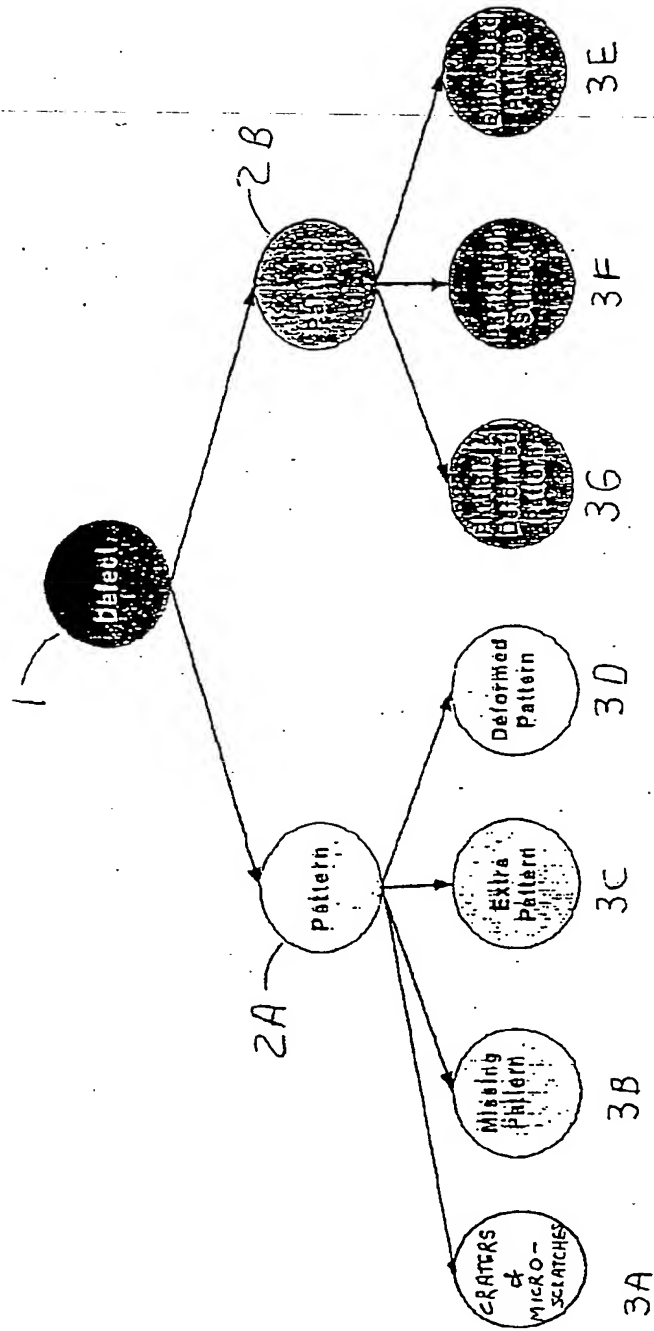
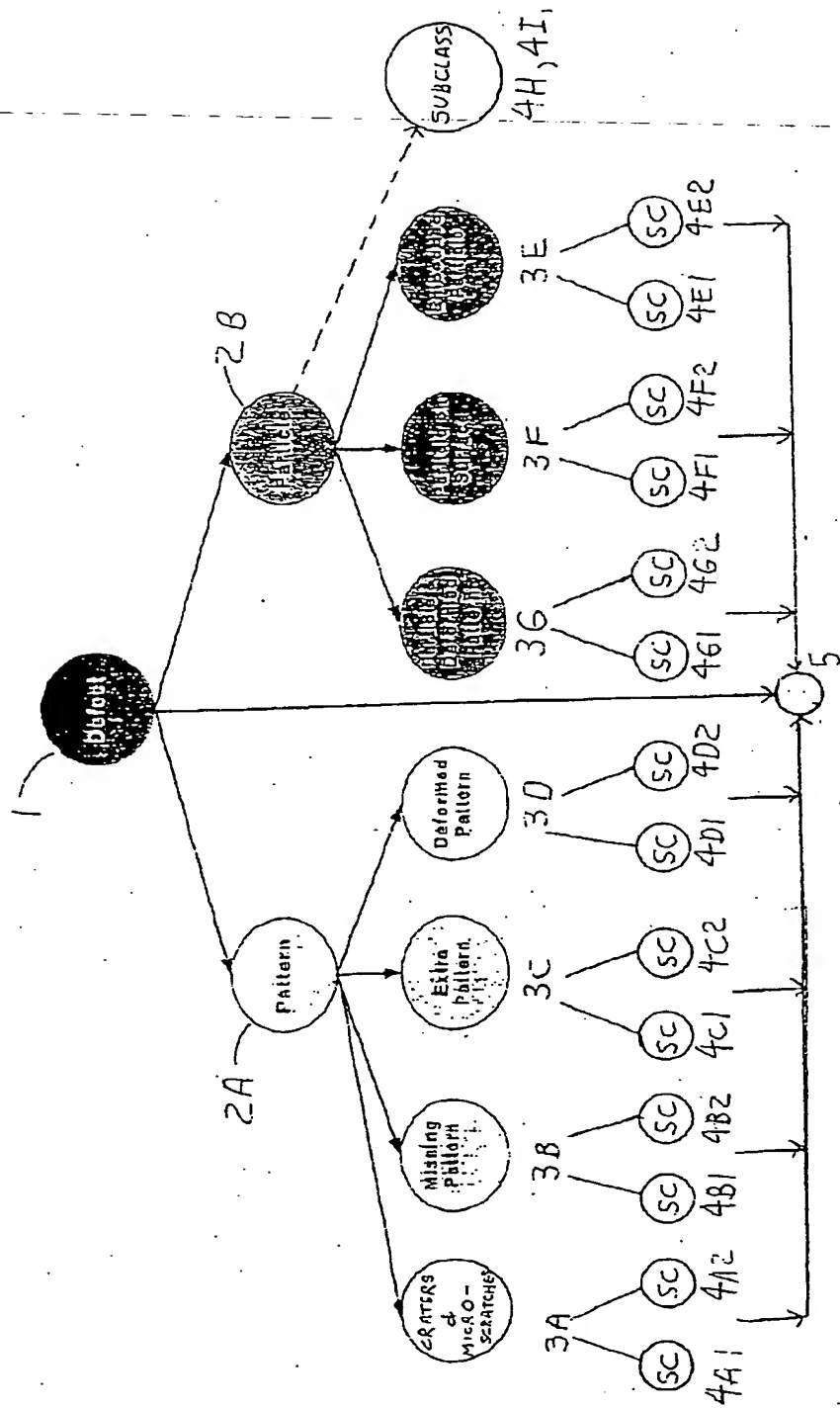
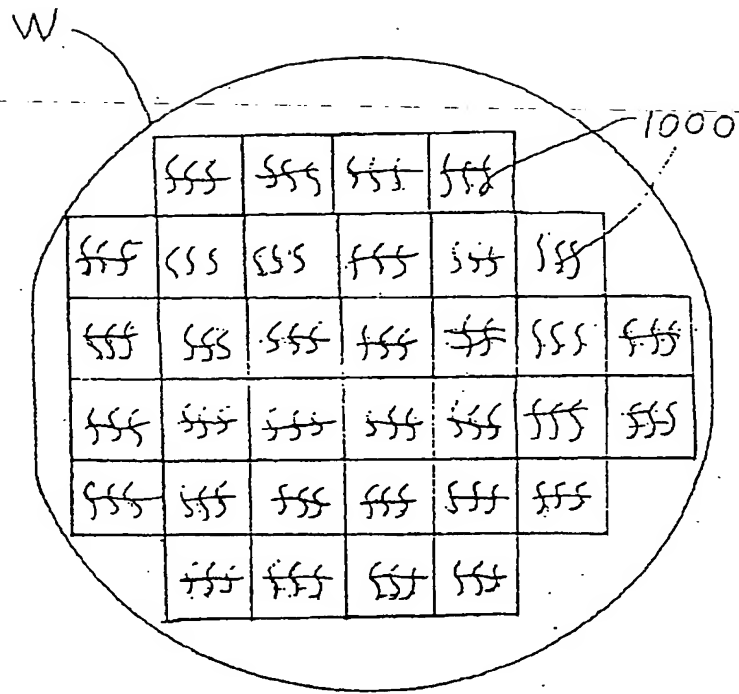


FIG. 1

FIG. 2





F16.3

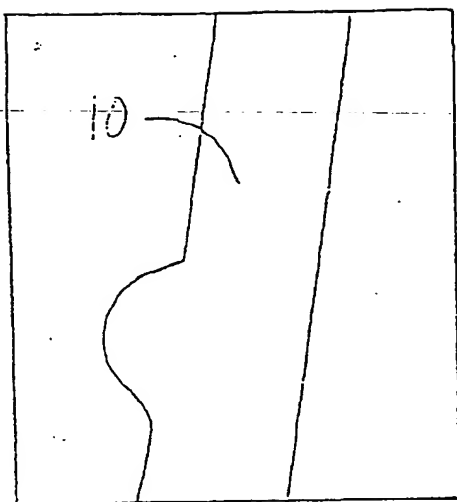


FIG. 4A

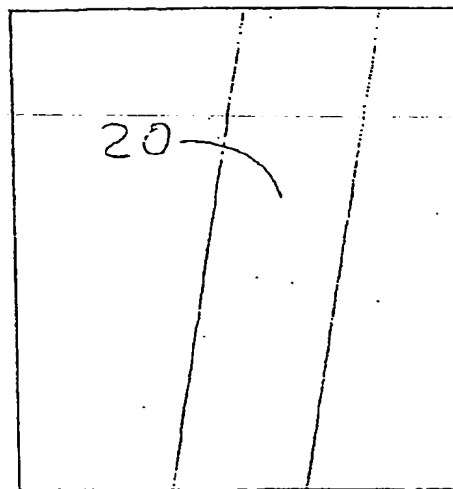


FIG. 4B

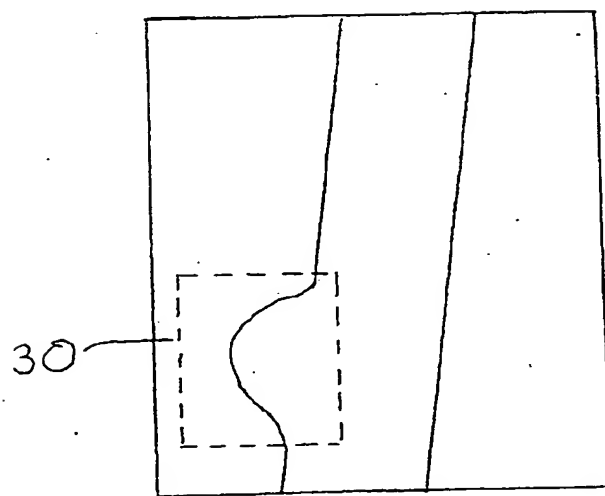


FIG. 4C

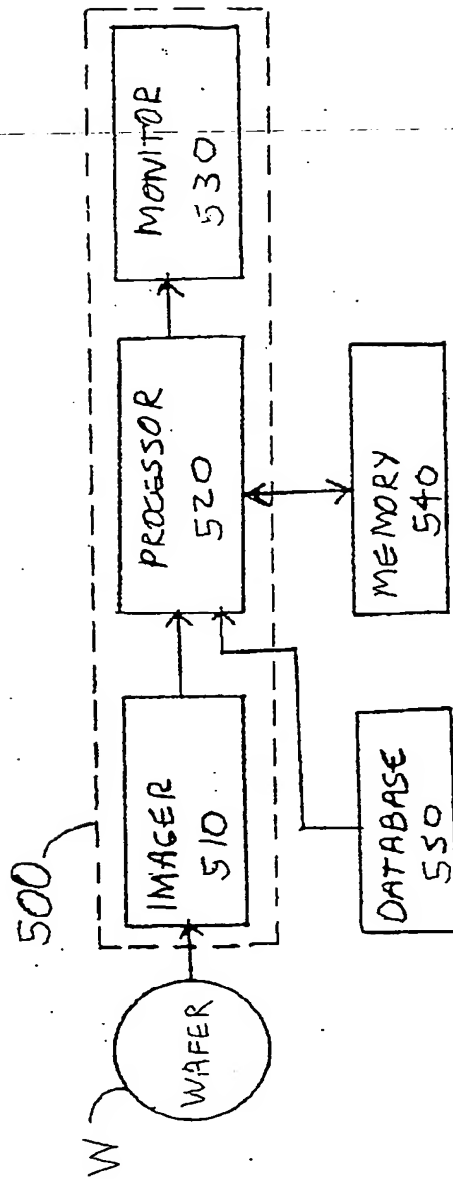


FIG. 5

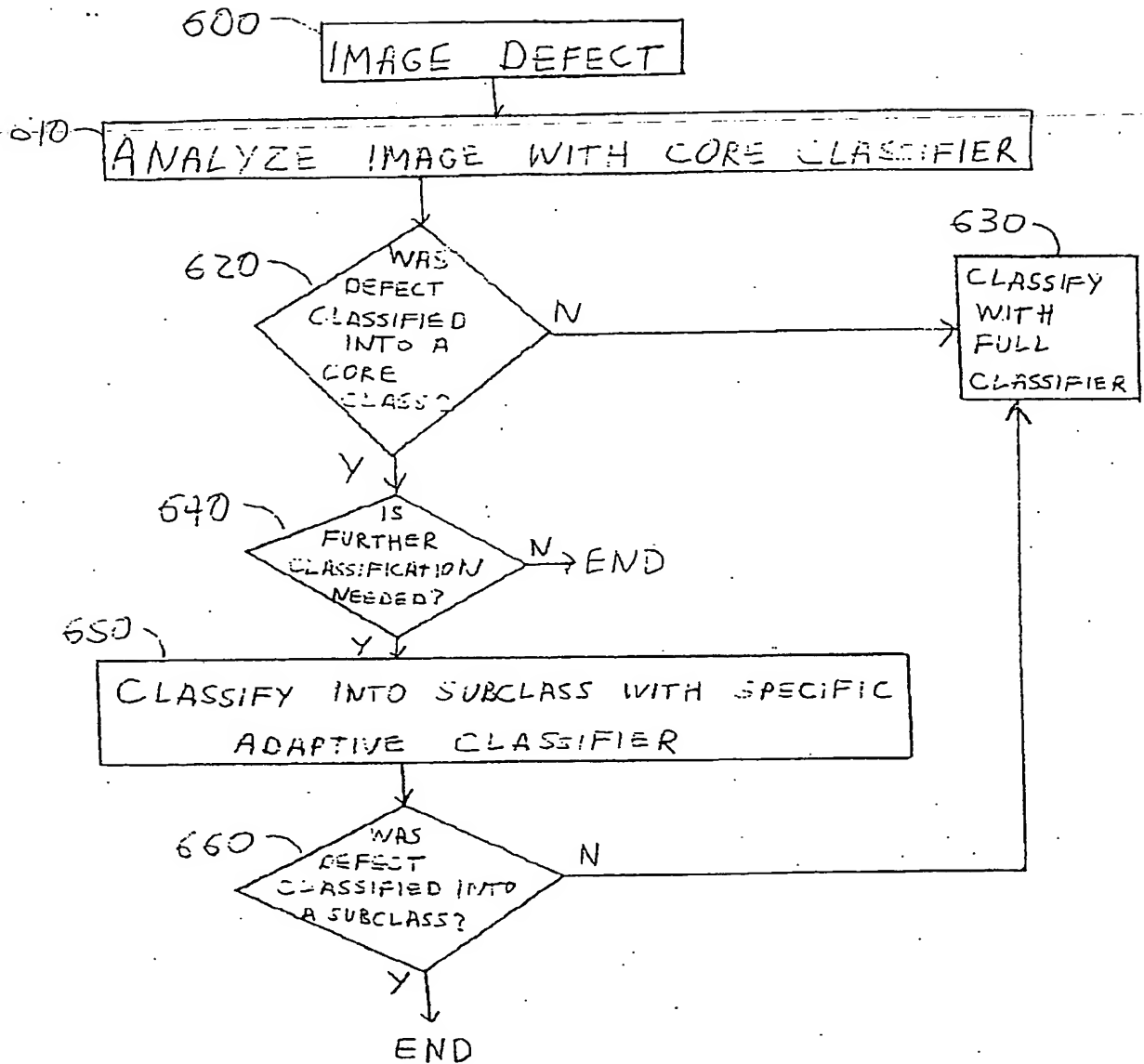


FIG. 6

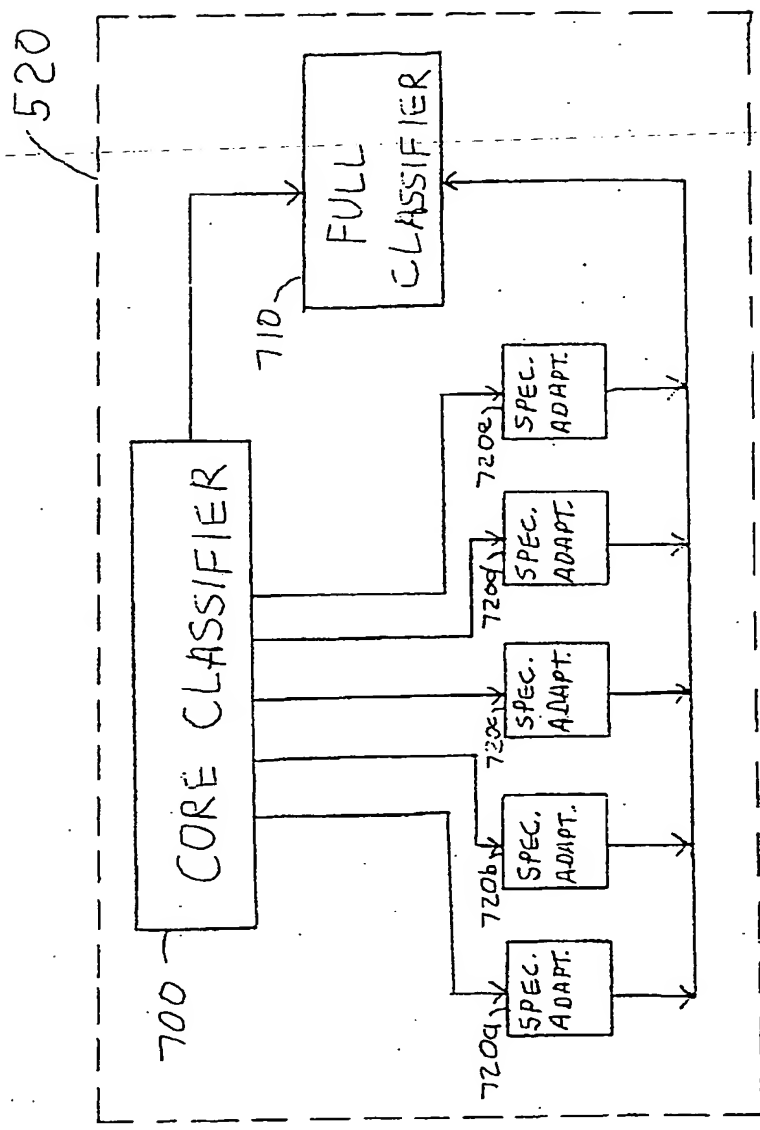
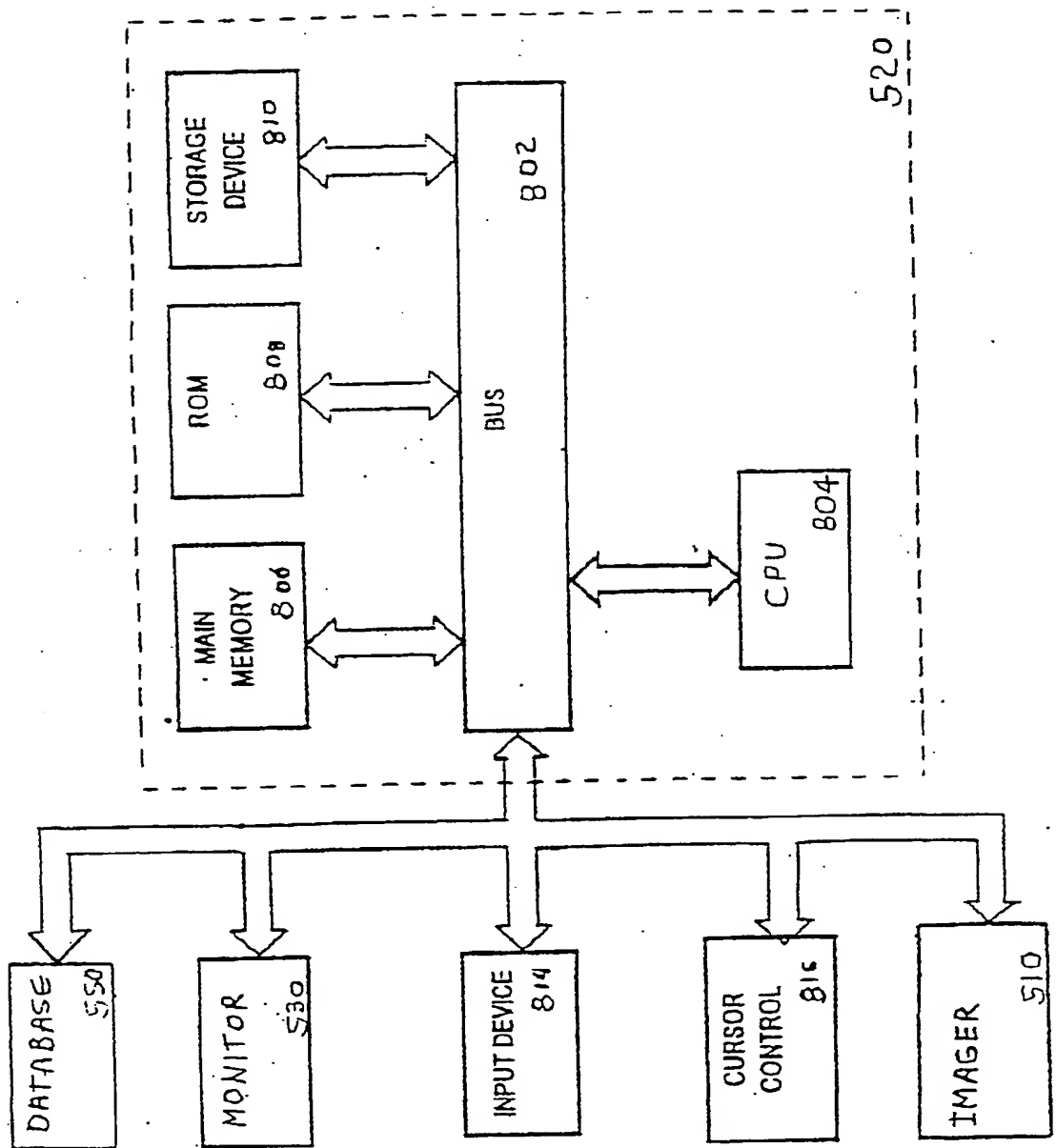


FIG. 7

FIG. 8



Abstract of the Disclosure

A method and apparatus is provided for automatically classifying a defect on the surface of a semiconductor wafer into one of a predetermined number of core classes using a core classifier employing boundary and topographical information. The defect is then further classified into a subclass of arbitrarily defined defects defined by the user with a specific adaptive classifier associated with the one core class and trained to classify defects only from a limited number of related core classes. Defects that cannot be classified by the core classifier or the specific adaptive classifiers are classified by a full classifier.